



José Pedro Galamba Egídio Reis

Licenciado em Ciências da Engenharia Eletrotécnica
e de Computadores

Estudo do Desempenho de um protocolo Slotted ALOHA P-persistente numa rede de Rádio Cognitivo

Dissertação para obtenção do Grau de Mestre em
Engenharia Eletrotécnica e de Computadores

Orientador: Luís Filipe Lourenço Bernardo, Professor Auxiliar, FCT-UNL

Júri:

Presidente: Doutor Rodolfo Alexandre Duarte Oliveira – FCT-UNL

Arguentes: Doutora Teresa Maria Sá Ferreira Vazão Vasques –
IST/UL

Doutor Luís Filipe Lourenço Bernardo – FCT-UNL



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Março 24, 2015

Estudo do Desempenho de um protocolo *Slotted ALOHA P*-persistente numa rede de Rádio Cognitivo

Copyright © José Pedro Galamba Egídio Reis, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Agradecimentos

Este espaço é dedicado àqueles que contribuíram, direta ou indiretamente, para que esta dissertação fosse realizada.

Em primeiro lugar, gostaria de agradecer ao orientador desta dissertação. Ao Professor Dr. Luís Filipe Bernardo pela sua colaboração, conhecimentos transmitidos e dedicação notável no acompanhamento deste estudo. Agradeço também as suas verificações, reconsiderações e revisões prestadas.

Agradeço também ao Professor Dr. Rodolfo Alexandre Oliveira, pela disponibilidade e auxílio ocasional ao longo de todo o trabalho.

Agradeço ainda o apoio fornecido pela Fundação para a Ciência e Tecnologia, através da bolsa de investigação no âmbito do Projeto *PTDC/EEA-TEL/120666/2010*.

Aos meus amigos do curso pela paciência e apoio prestados ao longo de todo o percurso académico.

Aos meus amigos de infância pelos momentos de desabafo e palavras de animo quando a motivação parecia desaparecer.

E porque o melhor fica sempre para o final, agradeço aos meus pais pelo apoio incondicional durante todo o processo de dissertação, por todos os valores e educação que me transmitiram e pelo incentivo de estudar e terminar o ensino superior.

A realização desta dissertação marca o fim de uma importante etapa da minha vida. A todos eles, deixo o meu sincero agradecimento.

Resumo

O surgimento de produtos e serviços sem fios levou à necessidade de maiores débitos. Consequentemente, a procura pelo espectro eletromagnético atingiu valores nunca antes vistos, tornando-o escasso e com um enorme valor comercial. Um grande paradigma da atualidade para os investigadores é como fazer a gestão deste espectro de forma a aumentar a capacidade e a utilização mais eficiente deste através de novas abordagens. O radio cognitivo surge como uma das abordagens inovadoras. Estas redes permitem que utilizadores secundários explorem o espectro não utilizado de tal forma que a interferência sobre os utilizadores primários do espectro seja mínima ou nula, enquanto que estes últimos mantêm os direitos legais de utilizar o seu espectro sempre que quiserem.

Este trabalho propõe um novo modelo de análise de desempenho de um protocolo *Slotted ALOHA P*-persistente de acesso ao meio numa rede de rádio cognitivo composto por um único utilizador primário e vários utilizadores secundários com processos de chegada de tráfego de *Poisson*. Os secundários utilizam um modelo de deteção baseada em energia, uma vez que a alteração do estado de atividade do primário é independente dos outros utilizadores. Os ciclos de operação dos secundários estão sincronizados e, quando estes têm um pacote para transmitir e o meio está livre, enviam-no com probabilidade P . Ao contrário de outros modelos propostos, esta dissertação propõe um novo modelo analítico para o atraso e débito dos secundários que considera a duração do estado de atividade do utilizador primário no atraso. As expressões do modelo são validadas através de um simulador cujos valores são obtidos sob várias condições. Esta tese estuda o impacto da otimização do valor P para duas diferentes configurações de rede: o valor P ótimo quando é conhecido o número de secundários e quando o número e carga são conhecidos.

Palavras Chave: Rádio cognitivo; *Slotted ALOHA P*-persistente motivação;

Avaliação de desempenho analítico; Otimização da rede secundária.

Abstract

The overgrowth of wireless networks and devices have led to an increased demand for spectrum. Given the limited available frequency spectrum, the current management policies cannot accommodate this increasing number of new networks and devices. Therefore, innovative approaches are being investigated. One of these approaches is the Cognitive Radio networks. These networks let opportunistic users exploit unused spectrum in such a way that they do not cause interference to the licensed users, while assuring that the latter users keep the legacy rights on the usage of their spectrum.

In this work, we propose a new performance analytical model of a P-persistent Slotted ALOHA medium access protocol in a radio cognitive network composed by a single primary user's transmitter and several secondary users with Poisson traffic. An energy-based sensing model is considered, since primary user changes his activity state independently of the secondary users. The last ones run a synchronized sensing-access operation cycle and when they have a packet to transmit and the channel is sensed free, they access the data slots with probability P . Unlike the other existing models, we propose a new analytical model for the secondary users' delay and throughput which considers the effect of the primary user's activity state duration. In addition, all the model expressions are validated by a simulator whose results are obtained under different network scenarios. This dissertation also studies the impact of the P value and proposes an optimization for two different context configurations: the optimal P value when only the number of secondary users is known, and when the number and load are known.

Keywords: Cognitive radio; P-persistent Slotted ALOHA; Analytical performance evaluation; Secondary network optimization.

Glossário

AP Access Point

AWGN Additive White Gaussian Noise

CDMA Code-Division Multiple Access

CR Cognitive Radio

CR-ALOHA Cognitive Radio Slotted ALOHA

CR-CSMA Cognitive Radio CSMA

CSMA Carrier Sense Multiple Access

FCFS First-Come First-Served

GDB Geolocation Data Base

IEEE Institute of Electrical and Electronics Engineers

ISM Industrial, Scientific and Medical

MAC Medium Access Control

MPR Multi Packet Reception

PU Primary User

PSD Power Spectral Density

RC Rádio Cognitivo

SU Secondary User

SNR Signal-to-Noise Ratio

SPR Single Packet Reception

TDMA Time-Division Multiple Access

WLAN Wireless Local Area Network

UHF Ultra High Frequency

VHF Very High Frequency

Conteúdo

Agradecimentos	iii
Resumo	v
Abstract	vii
Glossário	ix
1 Introdução	1
1.1 Introdução	1
1.2 Motivação	3
1.3 Objetivos e Contribuições	3
1.3.1 Objetivos	3
1.3.2 Contribuições	4
1.4 Estrutura da Dissertação	4
2 Trabalho relacionado	7
2.1 <i>Spectrum sensing</i> em Rádio Cognitivo	7
2.1.1 Dimensões de Sensing no Rádio Cognitivo	8
2.1.2 Métodos de deteção da ocupação do meio	10
2.1.3 Métodos de sensing para Rádio Cognitivo	11
2.2 Protocolos <i>MAC</i> em Rádio Cognitivo	16
2.2.1 Protocolos centralizados vs protocolos distribuídos	17
2.2.2 Métodos de acesso	19
2.3 Protocolos <i>MAC</i> de Rádio Cognitivo com canal único	20
2.3.1 Protocolo Slotted Cognitive Radio ALOHA (CR-ALOHA)	20
2.3.2 Protocolo <i>Cognitive Radio CSMA</i> (CR-CSMA)	23
2.3.3 Protocolo <i>Slotted ALOHA P</i> -persistente	25
2.3.4 Comparação entre os protocolos	27
3 Descrição do modelo	29
3.1 Descrição do Sistema	29
3.1.1 Comportamento do utilizador primário	30

3.2	Modelação do sistema	34
3.2.1	Primeira abordagem - Modelação <i>slot a slot</i>	34
3.2.2	Segunda abordagem - Modelação de tempo de PU	40
3.2.3	Terceira abordagem - modelação de tempo de PU e de erros	48
3.3	Otimização do Modelo	55
4	Desempenho do sistema	59
4.1	Implementação do simulador	59
4.1.1	Modelo geral do simulador	59
4.1.2	Parâmetros iniciais do sistema	60
4.1.3	Geração aleatória da chegada de pacotes	61
4.1.4	Controlador do avanço do tempo	62
4.1.5	Resultados da simulação	63
4.2	Análise da variação de P	64
4.2.1	Análise de P_{QE}	65
4.2.2	Análise do débito	65
4.2.3	Análise do atraso médio de pacotes	67
4.3	Análise do desempenho	69
4.3.1	Escalabilidade do sistema	69
4.3.2	Impacto do P_U no sistema	71
5	Conclusões	75
5.1	Considerações Finais	75
5.2	Trabalho Futuro	76
	Anexos	77
	A Publicações	79
	Bibliografia	86

Lista de Figuras

2.1	Oportunidades no domínio da frequência e tempo [YA09].	8
2.2	Oportunidades no domínio do espaço geográfico [YA09].	9
2.3	Oportunidades no domínio dos códigos.	10
2.4	Oportunidades no domínio do ângulo de transmissão.	10
2.5	Comparação entre métodos de sensing [YA09].	16
2.6	Classificação dos protocolos CR-MAC [CC09].	17
2.7	Modelo da rede CR de [CLMW11].	20
2.8	Estrutura da trama MAC do <i>Slotted CR-ALOHA</i>	21
2.9	Topologia da rede do <i>Slotted CR-ALOHA</i> [HYS12].	26
2.10	Slotted ALOHA para os <i>PU</i> s [HYS12].	26
3.1	Trama de um <i>SU</i>	30
3.2	Cadeia de Markov do comportamento do <i>PU</i>	31
3.3	Dois cenários possíveis na alteração de estados do <i>PU</i>	32
3.4	Valor instantâneo e valor médio do número de pacotes na fila de espera ao longo do tempo.	37
3.5	Resultados da probabilidade da fila de espera estar vazia obtidos com o modelo e com o simulador.	39
3.6	Débito por <i>SU</i> obtido com o modelo e com o simulador.	39
3.7	Agregação de <i>slots</i> com <i>PU</i> ativo na segunda abordagem.	40
3.8	Slots com <i>PU</i> ativo e inativo. Representação de τ_0 e τ_1	43
3.9	Resultados do débito obtidos com o modelo e com o simulador.	47
3.10	Resultados da probabilidade da fila de espera estar vazia obtidos com o modelo e com o simulador.	47
3.11	Resultados do atraso médio por pacote obtidos com o modelo e com o simulador.	48
3.12	Resultados da probabilidade da fila de espera estar vazia obtidos do modelo e do simulador em função de P	49
3.13	P_{QE} em função da carga para o modelo e o simulador.	54
3.14	Resultados do débito obtidos do modelo e do simulador.	54
3.15	Resultados do atraso médio de um pacote obtidos do modelo e do simulador.	55

3.16	Resultados do débito máximo do sistema em função de J do modelo.	56
3.17	Débito médio medido com o modelo em função de P para $J\lambda < S^*$	57
3.18	Atraso médio de um pacote medido pelo modelo em função de P para $J\lambda < S^*$	58
4.1	Módulos do simulador.	60
4.2	Geração de pacotes.	62
4.3	Exemplo do controlador de tempo e ilustração dos instantes em que ocorrem os eventos.	64
4.4	P_{QE} em função de P . Impacto da variação de λ nos valores de P_{QE} para $J = 2$	65
4.5	P_{QE} em função de P . Impacto da saturação do sistema nos valores de P_{QE} . Resultados vermelho e verde referentes a $J = 3$. Os restantes foram simulados para $J = 2$	66
4.6	Débito em função de P . Impacto da variação de λ nos valores do débito para $J = 2$	66
4.7	Débito em função de P . Impacto da saturação do sistema nos valores do débito.	67
4.8	Atraso médio em função de P . Impacto da variação de λ nos valores do atraso médio para $J = 2$	68
4.9	Atraso em função de P . Impacto da saturação do sistema nos valores do atraso médio.	68
4.10	P_{QE} em função de λJ . Impacto da variação de J nos valores de P_{QE}	69
4.11	Débito em função de $J\lambda$. Impacto da variação de J nos valores do débito.	70
4.12	Atraso médio em função de λJ . Impacto da variação de J nos valores do atraso médio.	71
4.13	P_{QE} em função de λ . Impacto da variação de P_U nos valores de P_{QE}	72
4.14	Débito em função de λ . Impacto da variação de P_U nos valores do débito.	72
4.15	Atraso médio em função de λ . Impacto da variação de P_U nos valores do atraso médio.	73

Lista de Tabelas

2.1	Comparação entre os três protocolos.	28
3.1	Parâmetros de configuração da 1 ^a abordagem.	38
3.2	Parâmetros de configuração da 2 ^a abordagem.	46
3.3	Parâmetros de configuração da 3 ^a abordagem.	53
3.4	Síntese das abordagens.	55
3.5	Parâmetros de configuração da figura 3.16.	56
3.6	Parâmetros de configuração da figura 3.17 e 3.18.	57
4.1	Parametrização de configuração da variação de P	64
4.2	Parametrização de configuração da escalabilidade do sistema.	69
4.3	Parametrização de configuração do impacto do P_U no sistema.	71

Capítulo 1

Introdução

1.1 Introdução

Ao longo dos últimos anos, o mundo presenciou um crescimento acelerado do número de redes e aparelhos sem fios. Atualmente, quase todas as casas, locais públicos ou instalações governamentais possuem redes sem fios que têm a capacidade de trabalhar com vários aparelhos ao mesmo tempo. Estas redes possibilitam a transferência de dados na rede local e na Internet, assim como o acesso a outros serviços e, portanto, têm um papel importantíssimo na nossa sociedade [DSB12]. Este aumento de utilização e a necessidade de maiores débitos levou ao aumento da utilização de espectro eletromagnético, tornando-o escasso e com um enorme valor comercial.

Enquanto os gráficos de alocação de frequência divulgam que praticamente todas as bandas de frequência já estão atribuídas, estudos revelam que as estratégias de alocação estática tradicional de espectro provocam buracos temporais e geográficos [DSB12] no uso do espectro nas bandas licenciadas. Além disso, nestes últimos anos, as bandas industriais, científicas e médicas (*ISM*) não licenciadas têm permitido o desenvolvimento de tecnologias como o *Wireless Fidelity* (WiFi), o Bluetooth, telefones sem fios, etc. O grande sucesso destas bandas deu origem ao problema da coexistência de sistemas heterogêneos que podem interferir entre si.

O Rádio Cognitivo (*RC*) surge como uma forma de melhorar o acesso e a gestão de espectro tanto nas bandas licenciadas como nas bandas não licenciadas. Esta tecnologia procura oferecer uma chance de acesso aos utilizadores oportunistas nas bandas licenciadas

do espectro, sem que estes interfiram com a comunicação dos utilizadores licenciados.

Através das funções do *RC*, as redes de utilizadores secundários realizam uma monitorização do meio, também conhecida como *sensing*, para identificar qualquer oportunidade de espectro em que podem transmitir, de forma a assegurar o nível de proteção exigido pelos utilizadores licenciados. A partir da utilização do *sensing*, as redes de rádio cognitivo conseguem obter uma amostra do meio que as rodeia e, com base nesta amostra, utilizar espectro que seria desperdiçado para proveito próprio.

Atualmente, existem diversos protocolos de acesso ao meio propostos já adaptados a redes de rádio cognitivo. Por norma, é necessário efetuar o *sensing* do meio para verificar que não se afeta a transmissão do utilizador licenciado, antes de qualquer utilizador oportunista aceder ao meio. Visto que a maioria dos nós das redes cognitivas não tem a capacidade de fazer o *sensing* e transmitir ao mesmo tempo, grande parte dos protocolos de acesso ao meio realizam um *sensing* periódico para manterem a informação sobre o estado do meio atualizada. Desta forma, quando um utilizador licenciado inicia uma transmissão, os utilizadores oportunistas têm a possibilidade de evacuar o canal rapidamente.

Este trabalho propõe um modelo analítico alternativo para o desempenho de um dos diversos protocolos já existentes para redes cognitivas: o protocolo *Slotted ALOHA P*-persistente. Esta rede é constituída por um utilizador licenciado e vários oportunistas. Os últimos contêm uma fila de espera onde podem armazenar pacotes para que, caso o meio esteja ocupado, possam transmiti-los mais tarde, assim que este estiver livre. Como tal, é modelado o débito e o atraso médio de pacotes de um utilizador oportunista, considerando que este utiliza uma abordagem de *sensing* baseado em energia. Além disso, o modelo contém uma secção de otimização do valor de acesso ao meio *P*, com o objetivo de maximizar o aproveitamento do meio não utilizado pela rede primária, sem que esta seja prejudicada. Por fim, os resultados obtidos através da modelação da rede são comparados aos de um simulador com o intuito de os validar.

Como a evolução da tecnologia tem a tendência natural de exigir cada vez mais eficiência na utilização do espectro para poder acompanhar o aumento do número de aparelhos, é provável que as redes de rádio cognitivo sejam do interesse dos sectores industrial e académico devido às suas potencialidades [CC09].

1.2 Motivação

As características de propagação das bandas de operação da televisão são muito desejáveis e convenientes para muitos serviços de transmissão sem fios. Estas bandas residem abaixo da frequência de 1 GHz, onde a obstrução material é menos prejudicial em relação às altas frequências, permitindo cobertura sem linha de vista e apresentando vantagens de *path loss* sobre as bandas *ISM*. A ocupação do espectro eletromagnético reduziu-se com a transição da televisão analógica para a digital, vagando certas bandas na gama *very high frequency* (*VHF*) e *ultra high frequency* (*UHF*). Consequentemente, surgiram duas normas de rádio cognitivo que aproveitam este espectro: o *IEEE* 802.11af [FGK⁺13b] e o *IEEE* 802.22 [SCL⁺09b].

O *IEEE* 802.22 é um protocolo baseado em rádio cognitivo que tem o objetivo de levar às zonas rurais o acesso a banda larga. Esta norma tem um alcance de 17 a 30 km desde a estação base até ao utilizador [SCL⁺09b]. A norma 802.11af tem o objetivo de aumentar o alcance do wi-fi até 5km, sem a utilização de mais intra-estruturas, através da utilização das bandas de TV e de técnicas de rádio cognitivo [FGK13a]. As duas normas utilizam uma entidade central que coordena a rede dos utilizadores oportunistas com a ajuda de uma estrutura de dados. Esta estrutura ou base de dados contém informação da atividade dos utilizadores licenciados e dados e parâmetros de transmissão necessários para aceder à rede.

Para além das técnicas baseadas em bases de dados, também é possível definir outros protocolos baseados unicamente em técnicas de monitorização local (*sensing*) do canal. Entre elas, pode-se considerar a utilização de um protocolo *Slotted-Aloha P*-persistente para terminais com um único rádio para coordenar o acesso ao espetro deixado livre pelos utilizadores licenciados.

1.3 Objetivos e Contribuições

1.3.1 Objetivos

Esta dissertação tem como objetivo desenvolver um modelo de desempenho para uma rede de rádio cognitivo com utilizadores licenciados e oportunistas.

O objetivo inclui o desenvolvimento de um modelo analítico para o desempenho do protocolo *Slotted Aloha P*-persistente numa rede cognitiva, que considere o desempenho do mecanismo de *sensing*. Também inclui o desenvolvimento de um simulador, que permita avaliar o desempenho do sistema e validar a precisão do modelo. Com os resultados do simulador, pretende-se compará-los aos do modelo analítico e analisá-los face à probabilidade da fila de espera estar vazia, ao débito do sistema e ao atraso médio por pacote. Numa última fase, procura-se a configuração ótima para o sistema. O objetivo é encontrar o valor de P ótimo que maximiza o débito do sistema, ou para uma carga fixa, que minimiza o atraso médio por pacote.

1.3.2 Contribuições

Esta dissertação analisa a utilização de um protocolo *Slotted Aloha P*-persistente numa rede cognitiva. As principais contribuições são:

- Um modelo analítico de desempenho para o sistema que modela o débito e atraso;
- Um simulador do sistema que inclui a modelação do *sensing* e da dinâmica dos nós;
- A determinação da configuração ótima do sistema para maximizar o débito ou para minimizar o atraso quando não está saturado.

Por fim, os resultados do trabalho contribuíram para uma publicação no *workshop IEEE SCAN* 2015. O artigo referido encontra-se incluído no Apêndice A.

1.4 Estrutura da Dissertação

A dissertação encontra-se dividida em cinco capítulos e um apêndice. No capítulo 2 são apresentados alguns conceitos já desenvolvidos na área de rádio cognitivo. É feito um levantamento geral de métodos para a deteção de espectro e aborda-se o tema dos protocolos de controlo de acesso ao meio (MAC) nas redes de rádio cognitivo. São apresentados também três sistemas de rádio cognitivo utilizando protocolos com características semelhantes ao protocolo analisado.

A terceira parte desta dissertação descreve detalhadamente o sistema proposto, bem como o comportamento do utilizador primário. Adicionalmente, são apresentados dois

modelos intermédios e o modelo final proposto, discutindo-se a validade das aproximações consideradas em cada um. No final do capítulo estuda-se a otimização do desempenho do sistema.

O capítulo 4 descreve de uma forma detalhada o simulador utilizado para validar o modelo teórico. Este capítulo é dividido em várias subsecções. Cada secção caracteriza um modulo do simulador. Além disso, este capítulo também contém a análise dos resultados obtidos através do modelo teórico e a comparação entre estes resultados e os valores das simulações. Esta parte analisa os valores ótimos do modelo em função ao débito, ao atraso médio por pacote e à probabilidade das filas dos utilizadores oportunistas estarem vazias, assim como investiga o comportamento do sistema face à escalabilidade e à atividade do utilizador licenciado. No final, estes resultados são comparados com os modelos descritos no capítulo 2.

O capítulo 5 resume as conclusões obtidas neste trabalho, apontando algumas linhas de investigação para trabalho futuro.

Por fim, o apêndice A contém o artigo publicado no Workshop mencionado nas Contribuições.

Capítulo 2

Trabalho relacionado

As redes de *RC* são redes dinamicamente reconfiguráveis que se adaptam ao ambiente envolvente, permitindo que os seus aparelhos possam explorar oportunisticamente as porções de espectro que os utilizadores de outras redes não utilizam. Estes últimos utilizadores são licenciados, isto é, têm permissão para transmitir em certas bandas de frequência. Os utilizadores licenciados também podem ser denominados de Utilizadores Primários (*PU*) e não estão cientes dos comportamentos cognitivos. Por outro lado, os utilizadores oportunistas são chamados de Utilizadores Secundários (*SU*) e são tipicamente não licenciados, portanto têm a responsabilidade de não interferir com as transmissões dos *PU*s visto que estes últimos têm uma prioridade maior de acesso ao meio. É importante referir que os *SUs* têm de ser capazes de sentir o meio envolvente, de modo a detetar ou não a ocupação do espectro.

Este capítulo introduz as duas principais funções do RC: a deteção de transmissões e o acesso ao meio. São apresentados os mecanismos de *sensing* mais conhecidos a nível físico assim como os protocolos de controlo de acesso ao meio (*MAC*) para redes de *RC* e as suas classificações. Por fim, são apresentados e comparados alguns protocolos *MAC* de rádio cognitivo.

2.1 *Spectrum sensing* em Rádio Cognitivo

A deteção do meio é um dos pontos mais importantes numa rede de rádio cognitivo, pois é desta forma que os *SUs* encontram as oportunidades de acesso.

2.1.1 Dimensões de Sensing no Rádio Cognitivo

A oportunidade de acesso decorre de uma banda de frequências que está desocupada, num determinado tempo e numa determinada área geográfica. Estas oportunidades são habitualmente exploradas em três dimensões do espectro eletromagnético: na frequência, no tempo e no espaço. Além destas três dimensões, também é possível obter oportunidades de exploração de espectro através de códigos ortogonais e ângulos/direção de emissão. No entanto, os algoritmos convencionais de deteção do meio não têm a capacidade de detetar sinais transmitidos pelos *PUs* através de espalhamento e saltos na frequência ou no tempo. Assim, ao analisar os códigos ortogonais, seria possível criar novos métodos de deteção de espectro com a capacidade de encontrar porções de espectro livre. Do mesmo modo, os sinais transmitidos através de antenas direcionais podem não ser detetados pelos algoritmos convencionais porque é assumido que os *PUs* e/ou os *SUs* transmitem em todas as direções. Por outro lado, com avanços recentes na tecnologia de multi-antenas, através da transmissão com *beamforming* (formação de feixe), tornou-se viável vários utilizadores comunicarem no mesmo canal, ao mesmo tempo e na mesma área geográfica [YA09]. Desta forma, surgiram novas oportunidades para os *SUs* utilizarem a banda sem interferir com os *PUs*.

Assim, é possível dividir o acesso ao meio nas várias dimensões mencionadas em cima. Estas vão ser analisadas de seguida [YA09]:

- **Frequência** - Os *SUs* exploram as oportunidades no domínio da frequência. As oportunidades nesta dimensão indicam que as bandas não estão todas ocupadas ao mesmo tempo e que algumas podem estar disponíveis para uso oportunista.

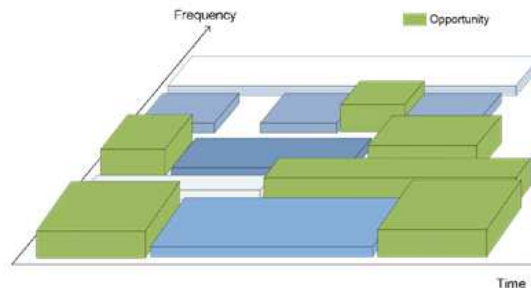


Figura 2.1: Oportunidades no domínio da frequência e tempo [YA09].

- **Tempo** - Na mesma banda, também há oportunidades de tempo, isto é, uma banda de frequência não é utilizada continuamente; podem existir espaços de tempo onde esta porção de banda estará disponível e poderá ser explorada.
- **Espaço geográfico** - Baseando-se na posição geográfica (latitude, longitude e altitude) e distância dos *PUs*, os utentes das redes de rádio cognitivo tiram proveito da atenuação do sinal por *path loss* (desvanecimento) para reutilizarem as bandas ocupadas pelos *PUs*. A ocupação do canal é determinada pelo nível de interferência no canal num determinado local e num determinado período. A ausência de interferência significa que não existe nenhuma transmissão de *PUs* numa determinada área.

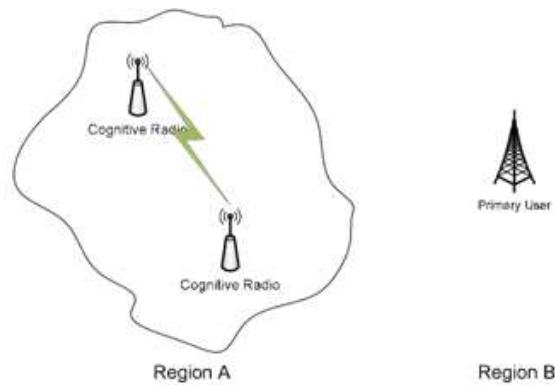


Figura 2.2: Oportunidades no domínio do espaço geográfico [YA09].

- **Códigos** - É possível realizar transmissões simultâneas sem interferir com as transmissões dos *PUs* utilizando códigos ortogonais. De realçar que é necessário existir uma sincronização entre os *SUs* e os *PUs*.
- **Ângulo de transmissão** - Para explorar as oportunidades nos ângulos de transmissão, os *SUs* têm que estar cientes da localização dos *PUs* assim como da direção do seu feixe de transmissão. Assim, os utentes cognitivos podem transmitir informação simultaneamente sem causar interferência aos *PUs*, visto que transmitirão para uma área diferente e com um ângulo diferente.

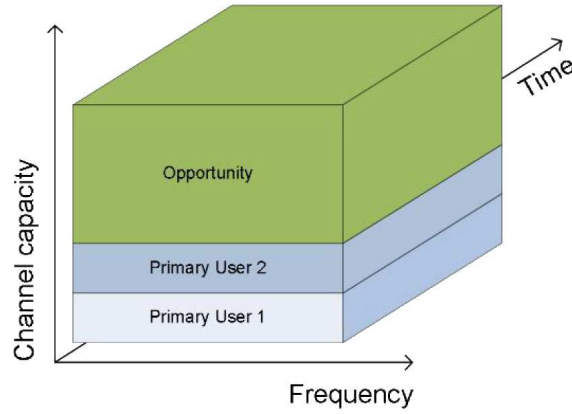


Figura 2.3: Oportunidades no domínio dos códigos.

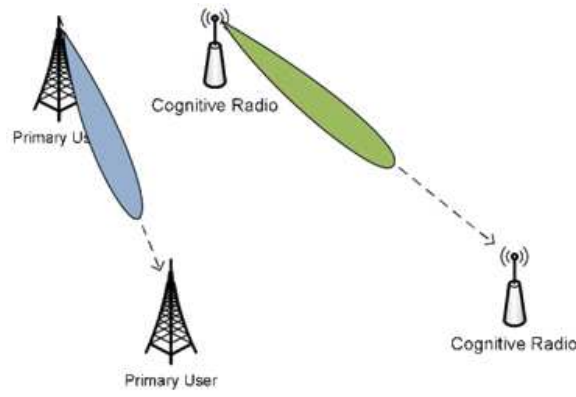


Figura 2.4: Oportunidades no domínio do ângulo de transmissão.

2.1.2 Métodos de detecção da ocupação do meio

É possível verificar a transmissão de um *PU* através da utilização de vários métodos que podem ser classificados em três tipos [SKMK10]:

- Detecção do sinal dos *PU*s ou *sensing* do meio;
- Detecção de *beacons* auxiliares;
- Utilização de bases de dados de geolocalização dos *PU*s (GDB).

O primeiro tipo de detecção da ocupação do meio é o que tem maior relevância para este documento porque a detecção do meio do modelo proposto é baseada neste método. Assim, o primeiro método é abordado com grande detalhe na secção 2.1.3.

No segundo tipo de detecção da atividade dos primários, são implementados transmissores de *beacons* na localização dos *PU*s. Estes transmissores emitem periodicamente *beacons* de notificação para avisar os *SUs* nas redondezas que não devem ocupar o canal.

Deste modo, garante-se aos *PU*s de baixa potência uma maior proteção face a interferências de *SUs* (problema do nó escondido, secção 2.2.2). No entanto, como os transmissores de *beacons* são infraestruturas fixas, não é viável serem implementadas em redes onde os *PU*s são móveis. Além disso, a alteração de todos os transmissores dos *PU*s para *beacons* auxiliares traria um custo excessivo.

Em alternativa à deteção do sinal dos primários ou *beacons* auxiliares, é possível utilizar bases de dados de geolocalização para reconhecer a presença de um *PU* num determinado canal. Assim, estas bases de dados têm o propósito de fornecer aos *SUs* os dados e parâmetros necessários para que possam ocupar os respetivos canais do meio sem que causem interferência aos primários. Esta informação contém os canais disponíveis para acesso oportunista, bem como o nível de potência do sinal de transmissão dos *SUs*. Assim sendo, é imprescindível que os *SUs* estejam equipados com GPS para que possam aceder às GDBs e que se saiba, *a priori*, a localização dos utilizadores primários [GBE⁺08].

2.1.3 Métodos de sensing para Rádio Cognitivo

Esta secção aborda o grupo de métodos de *sensing* ou deteção de espectro. Os métodos de deteção de espectro ainda estão na fase inicial de desenvolvimento, porém, já foi desenvolvida alguma diversificação na forma de identificar a presença de transmissões primárias. Em baixo, são apenas abordados alguns dos principais métodos de *sensing* do *RC*.

***Sensing* baseado na deteção de energia**

A abordagem de deteção de energia é o método mais comum de deteção de oportunidades de acesso por causa da sua baixa complexidade de implementação e computacional [YA09]. Além disso, os recetores têm a vantagem de não necessitarem de qualquer conhecimento do sinal dos *PU*s. Este método examina o nível de energia no meio durante um determinado período e compara-o com um determinado valor de limiar, γ , que depende do ruído no meio envolvente [Urk67]. A métrica para a decisão da deteção do nível de energia do meio pode ser escrita como

$$M = \sum_{n=0}^N |y(n)|^2. \quad (2.1)$$

A decisão de ocupação do espectro é obtida através da comparação do valor de M com o valor γ .

O sinal examinado pode resumir-se a duas hipóteses:

$$\mathcal{H}_0 : y(n) = w(n), \quad (2.2)$$

$$\mathcal{H}_1 : y(n) = s(n) + w(n), \quad (2.3)$$

onde $s(n)$ é o sinal detetado, $w(n)$ é o ruído branco Gaussiano (*AWGN*) detetado e n é o índice de amostras. De notar que se $s(n) = 0$, então o sinal detetado é constituído apenas por ruído, não existindo nenhuma transmissão de sinal de *PU*s no canal, ou seja, este cenário corresponde à hipótese \mathcal{H}_0 . Por outro lado, um cenário onde um *PU* está a utilizar o canal corresponde à hipótese \mathcal{H}_1 .

O desempenho deste algoritmo é afetado por duas probabilidades: a probabilidade de deteção P_D e a probabilidade de falso alarme P_{FA} . P_D é a probabilidade de deteção do sinal do *PU* quando esta transmissão está realmente a ser transmitida. Deste modo, é do interesse da rede cognitiva que o P_D seja o mais elevado possível. Esta probabilidade pode ser formulada da seguinte forma:

$$P_D = Pr(M > \gamma | \mathcal{H}_1), \quad (2.4)$$

ou seja, é a probabilidade da deteção do nível de energia ser maior que o limiar γ , quando existe um sinal de *PU* a ser transmitido no meio.

Por outro lado, P_{FA} é a probabilidade do recetor considerar que detetou a transmissão de um primário quando, na verdade, não existe nenhum *PU* a transmitir nessas frequências. P_{FA} é definido como

$$P_{FA} = Pr(M > \gamma | \mathcal{H}_0), \quad (2.5)$$

sucedendo quando o nível de energia do ruído é superior ao valor γ . De salientar que é vantajoso para a rede que o valor de P_{FA} seja o mais baixo possível, para prevenir a subutilização de oportunidades de transmissão.

O limiar de decisão γ é selecionado para se obter o melhor ajuste entre P_D e P_{FA} . Para isso, é necessário conhecer o ruído e a potência do sinal detetado. O ruído pode ser estimado mas a potência do sinal recebido é difícil de determinar, visto que este sofre alterações com as características das transmissões em curso e com a distância entre os utilizadores secundários e primários.

O ruído branco pode ser modelado através de uma variável aleatória Gaussiana com média nula e variância σ_w^2 ($w(n) = \mathcal{N}(0, \sigma_w^2)$). A modelação do sinal $s(n)$ também é feita através de variáveis aleatórias Gaussianas de média nula e variância σ_s^2 , cujo valor reflete o desvanecimento do sinal no meio. A variável aleatória M resulta da soma do quadrado de variáveis Gaussianas, logo tem uma distribuição chi-quadrado com grau de liberdade $2N$ e por isso, pode ser modelada como:

$$M = \begin{cases} \frac{\sigma_w^2}{2} \chi_{2N}^2 & \mathcal{H}_0, \\ \frac{\sigma_w^2 + \sigma_s^2}{2} \chi_{2N}^2 & \mathcal{H}_1 \end{cases} \quad (2.6)$$

Este método de deteção é considerado o mais importante para este documento, visto que tanto os protocolos analisados e comparados neste capítulo como o modelo da dissertação utilizam o método de sensing baseado na deteção de energia.

***Sensing* baseado na forma de onda**

É possível realizar a deteção de espectro através da correlação do sinal recebido com uma cópia conhecida (padrão conhecido) do próprio [YA09]. Este processo só é aplicável em redes com conhecimento de padrões de sinal e é chamado de *sensing* baseado na forma de onda.

Considerando o sinal $y(n)$ do método anterior, a métrica de decisão do *sensing* baseado na forma pode ser obtida através da seguinte expressão:

$$M = \text{Re} \left[\sum_{n=1}^N y(n) s^*(n) \right], \quad (2.7)$$

onde $y(n)$ é o sinal recebido e $s^*(n)$ representa o conjugado do sinal enviado (padrão conhecido).

Tal como no *sensing* baseado em deteção de energia, conclui-se se existe ou não transmissão de utilizadores primários através da comparação do valor da métrica com um valor de limiar.

***Sensing* baseado em ciclo estacionariedade**

As particularidades da ciclo estacionariedade são causadas pela periodicidade do sinal, pela média e pela auto-correlação e estas podem ser induzidas intencionalmente para auxiliar a deteção de espectro [Gar91] [MBA⁺07].

A deteção baseada em ciclo estacionariedade é uma técnica utilizada para detetar transmissões de utilizadores primários no meio envolvente através da exploração das características ciclo estacionárias do sinal recebido [LKHP07]. Assim, em vez de densidade espectral, a deteção do sinal num dado espectro utiliza a função de ciclo-correlação. Esta função consegue diferenciar o ruído branco do sinal transmitido pelo utilizador primário devido ao facto do ruído ser estacionário e, por isso, este terá correlação nula quando modulado com sinais ciclo estacionários [CB05]. Além disso, com esta técnica é possível distinguir diferentes transmissões originárias de utilizadores primários distintos.

***Sensing* baseado em identificação de rádio**

É possível obter um conhecimento total sobre as características do espectro através da identificação da tecnologia de transmissão utilizada pelos utilizadores primários. Tal conhecimento permite aos utilizadores de rádio cognitivo ter uma maior adaptação ao meio [YA06].

No *sensing* baseado em identificação de rádio, este conhecimento total é extraído do sinal recebido. Com as características retiradas, é possível descobrir a tecnologia utilizada do utilizador primário mais provável. As características obtidas a partir do *sensing* baseado em deteção de energia, por exemplo, são utilizadas para métodos de classificação [VFECH01] [MDV⁺01].

Assim, a finalidade desta técnica de *sensing* é identificar a presença de algumas técnicas de transmissão conhecidas e conseguir estabelecer uma comunicação através delas.

Sensing utilizando filtros adaptados

Esta técnica utiliza um filtro linear que maximiza o rácio entre a potência do sinal e a potência do ruído (SNR). A operação do filtro adaptado é equivalente a fazer a convolução entre o sinal desconhecido e um espelho do sinal de referência transladado no tempo [Tur04].

Este método é conhecido como a melhor técnica para detetar a transmissão de utilizadores primários quando é conhecido *a priori* o sinal transmitido. A sua principal vantagem é o curto período de tempo para atingir uma determinada probabilidade de falso alarme ou de deteção quando comparada com os outros métodos já abordados. No entanto, é exigido que os utilizadores de rádio cognitivo façam a desmodulação dos sinais recebidos. Consequentemente, é necessário um conhecimento perfeito sobre as características dos sinais dos utilizadores primários. Além disso, como os utilizadores secundários precisam de recetores para cada tipo de sinal recebido, a complexidade de implementação destes aparelhos de *sensing* torna-se pouco viável. Por fim, os filtros adaptados também consomem uma grande quantidade de potência visto que os recetores estão constantemente a executar vários algoritmos de deteção.

Comparação entre os métodos de *sensing*

Nesta secção são comparados os métodos de deteção de oportunidades de espectro abordadas até agora, desde a sua complexidade até à sua precisão de *sensing* (figura 3.3).

Quanto à robustez de métodos, o *sensing* baseado na forma de onda consegue ser mais robusto que as técnicas baseadas em deteção de energia e ciclo estacionariedade devido ao processamento coerente que surge do uso da componente determinística do sinal. No entanto, é necessário existir informação *a priori* das características dos utilizadores primários e estes devem transmitir padrões conhecidos.

Por outro lado, quando um utilizador cognitivo se depara com um cenário em que não existe a possibilidade de se conhecer as características dos utilizadores primários, o único método adequado é a técnica baseada em deteção de energia. Contudo, o ruído pode não ser estacionário e a sua distribuição pode não ser conhecida e, por isso, é possível que o ruído comprometa a decisão da ocupação do canal [CTB06]. Além disso, este

método também tem a desvantagem de sofrer de efeitos de filtros de banda base e de tons adulterados [MMB07].

Contrariamente à detecção de energia, o método baseado em ciclo estacionariedade tem um bom desempenho na presença de interferências de canais adjacentes ou co-canais, onde o ruído não é estacionário [TCB07]. Em contrapartida, o desvanecimento do canal e a vulnerabilidade aos *offsets* da amostragem do relógio criam uma diminuição do desempenho desta técnica.

Por fim, para a elaboração desta dissertação, optou-se pela utilização do *sensing* baseado em energia devido à sua simplicidade e facilidade de implementação.

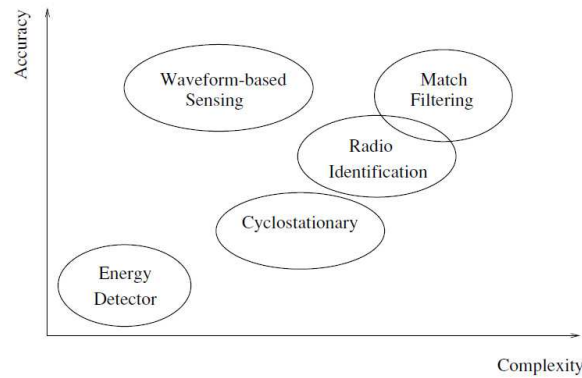


Figura 2.5: Comparação entre métodos de sensing [YA09].

2.2 Protocolos *MAC* em Rádio Cognitivo

O controlo de acesso ao meio tem uma grande importância para as funções de rádio cognitivo, principalmente na detecção de canais e partilha do espectro. Os protocolos *MAC* permitem que múltiplos utilizadores partilhem os recursos de espectro existentes determinando a ordem de acesso ao canal. Os protocolos *MAC* para rádio cognitivo (*CR-MAC*) surgiram inicialmente a partir de protocolos de redes *ad-hoc*. No entanto, os protocolos *CR-MAC* são mais complexos porque têm funcionalidades de detecção de espectro sofisticadas e requisitos de rápida adaptabilidade, pois têm que evacuar o meio quando o *PU* transmite. Os protocolos *CR-MAC* têm o objetivo de fornecer meios eficientes de detecção de canal para determinar a sua ocupação e partilhar o espectro entre outros utilizadores cognitivos da rede, desde que mantenham os níveis de interferência toleráveis para os

utilizadores prioritários [CC09].

Os protocolos *CR-MAC* podem ser divididos em dois grandes grupos: protocolos para redes de *RC* centralizadas e para redes de *RC* distribuídas (ou redes de *RC Ad Hoc*). Dentro de cada um destes grupos, os protocolos para redes de *RC* ainda podem ser classificadas como: protocolos de acesso aleatório, protocolos *time-slotted* ou protocolos híbridos.

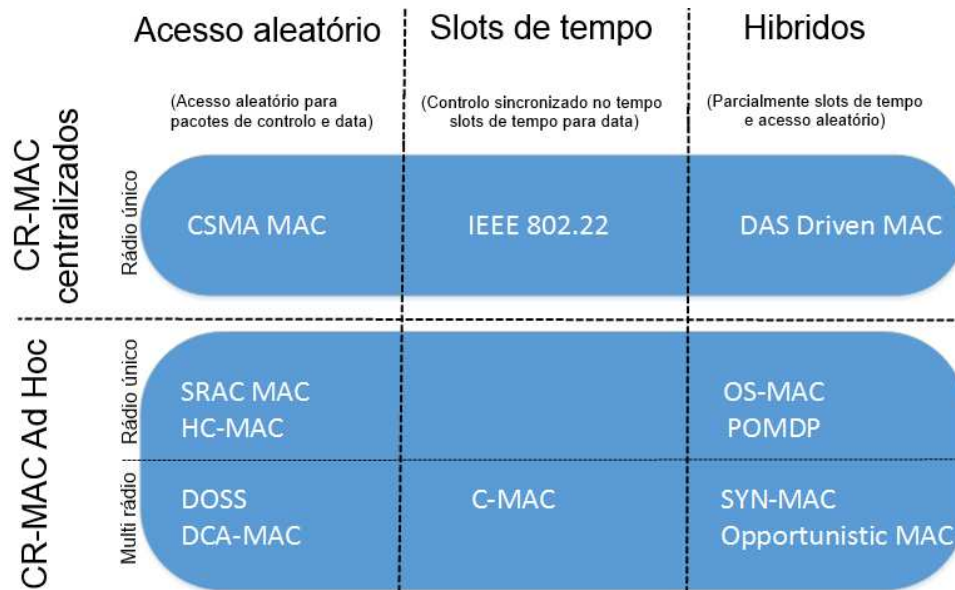


Figura 2.6: Classificação dos protocolos CR-MAC [CC09].

2.2.1 Protocolos centralizados vs protocolos distribuídos

Os protocolos centralizados precisam, obrigatoriamente, de uma entidade central, tal como uma estação base, que gere as atividades da rede, coordena as operações e sincroniza os nós. No entanto, a entidade central é estática e, geralmente, está à distância de um *hop* (salto) dos utilizadores secundários móveis. Os nós da rede enviam informação periodicamente sobre o seu estado atual (o tipo de dados que enviam é relativo e depende do protocolo utilizado). Numa topologia centralizada, o nível de potência de cada nó da rede é monitorizado pela entidade central com o propósito de controlar a potência de transmissão de todos os nós secundários [MSR07]. Desta forma, é possível reduzir a interferência entre nós da rede secundária e da rede primária e até reduzir a energia utilizada para a transmissão, aumentando assim o desempenho da rede secundária. Assim,

a abordagem centralizada no *RC* fornece um benefício de total conhecimento da rede, permitindo a possibilidade de desenvolver protocolos com uma eficiência mais elevada. [ZC08] sugere um protocolo *MAC* centralizado com uma utilização elevada de espectro para controlo, oferecendo uma abordagem livre de colisões e garantindo qualidade de serviço através da troca de informação entre os utilizadores secundários e a entidade central. No entanto, as redes centralizadas têm uma grande desvantagem: a informação precisa de ser sempre enviada para a entidade central, processada e propagada, novamente, para todos os nós da rede. Desta forma, o tráfego de controlo contribui para aumentar a carga no meio sem fios, existindo o risco de se atingir a saturação, especialmente em redes com muitos utilizadores; ou seja, os protocolos centralizados não são escaláveis.

Por outro lado, os protocolos distribuídos não possuem uma unidade central, pelo que têm uma organização e desenvolvimento mais flexíveis. Desta forma, os protocolos distribuídos têm a vantagem de reagir às mudanças do meio rapidamente. Embora sejam escaláveis, o *sensing*, a partilha de espectro e o acesso ao meio nestes protocolos são funções que precisam de um aumento de cooperação entre os vizinhos.

Os protocolos podem operar em sistemas de canal único ou multi-canal. Nos sistemas de canal único as redes primária e secundária coexistem na mesma banda de espectro. Estes protocolos têm a desvantagem do seu débito máximo obtido estar limitado. Por outro lado, num sistema multi-canal, a rede secundária tem a possibilidade de aceder ao meio através de um conjunto de canais. Este sistema tem a possibilidade de dispensar um canal comum apenas para controlo, que contribui para a interação e coordenação entre os SUs da rede.

Em [MSR07], é proposto um protocolo distribuído chamado *single radio adaptive channel MAC*. Neste protocolo, cada utilizador secundário guarda a frequência dos canais de receção dos seus vizinhos através de uma comunicação cooperativa. Assim, sempre que deseje comunicar com um nó vizinho, transmite a informação no canal de receção do respetivo recetor e recebe a resposta através de outra banda de frequência.

Para finalizar, é fundamental clarificar que o tráfego das redes com protocolos centralizados não tem que ser necessariamente maior do que para as redes distribuídas. Caso se pretenda atingir o mesmo nível de coerência da informação partilhada num protocolo

distribuído e num protocolo centralizado, são geralmente necessárias mais mensagens para o protocolo distribuído (e.g. capítulo 7 em [Tan03]).

2.2.2 Métodos de acesso

Foi referido anteriormente que, do ponto de vista de acesso, os protocolos *MAC* para *RC* podem ser classificados em: acesso aleatório, baseados em *slots* de tempo e híbridos.

Os protocolos *MAC* de *acesso aleatório* não precisam de sincronização entre os seus utilizadores. Sempre que algum utilizador secundário pretende transmitir, este monitoriza o meio primeiro. Se o canal não estiver ocupado, o utilizador acede ao meio.

Se, por outro lado, o canal estiver ocupado (tanto com transmissões secundárias como primárias), existem várias alternativas. O *CSMA MAC* 1-persistente é um exemplo de um protocolo onde o utilizador continua a monitorizar o meio até que este vague e transmite de imediato. Por norma, quando ocorrem colisões entre transmissões de utilizadores secundários, os *SUs* só voltam a monitorizar o canal e a transmitir após um período de tempo aleatório. Numa abordagem alternativa, como o *ALOHA*, espera um tempo aleatório antes de voltar a tentar.

Nos protocolos com acesso baseado em *slots* de tempo, a sincronização entre utilizadores é obrigatória. Para a rede secundária, o tempo é dividido em *slots* tanto para o período de *sensing* como para o período de transmissão. Ao contrário do que acontece com os protocolos de acesso aleatório, nos protocolos baseados em *slots* de tempo, todos os *SUs* da rede monitorizam o meio ao mesmo tempo. A norma 802.22 é um exemplo de um protocolo baseado em *slots* de tempo atualmente implementado [SCL⁺09a].

Nos protocolos híbridos, o período de transmissão é parcialmente dividido em *slots*, nos quais ocorre a sinalização de controlo. No entanto, o restante tempo de transmissão é acedido aleatoriamente, sem sincronização de tempo. O *Opportunistic MAC* [SZ12] é um exemplo de um protocolo híbrido que obriga os *SUs* a utilizarem dois *transceivers*, um para um canal comum de controlo e outro para aceder a uma banda de transmissão. O tempo é dividido em *slots* para a banda de espectro de transmissão, enquanto que o canal comum de controlo opera parcialmente em *slots* para a fase de informação, seguido de acesso aleatório para a fase de negociação.

2.3 Protocolos MAC de Rádio Cognitivo com canal único

Nesta secção são analisados e comparados três protocolos para redes de *RC* de canal único, ou seja, a rede primária e a rede secundária utilizam o mesmo canal de transmissão. Os protocolos analisados incluem o protocolo *Slotted Aloha p*-persistente mais dois protocolos comparáveis.

2.3.1 Protocolo Slotted Cognitive Radio ALOHA (CR-ALOHA)

Em [CLMW11], é apresentado um protocolo de acesso oportunista ao espectro baseado em Slotted ALOHA para uma rede de rádio cognitivo. Este protocolo é distribuído e o tempo é dividido em *slots*. É assumido que existe uma rede de *SUs*, que transmite num único canal partilhado com outra rede de *PU*s. A rede primária é constituída apenas por um transmissor e vários recetores e a rede secundária é formada por N *SUs* localizados dentro do alcance de transmissão da rede primária. Dentro do alcance da rede dos *PU*s, também se encontra um ponto de acesso secundário (*SAP*), capaz de comunicar com os *SUs* para resolver os seus problemas de sincronização. Na figura 2.7 é apresentado o modelo da rede.

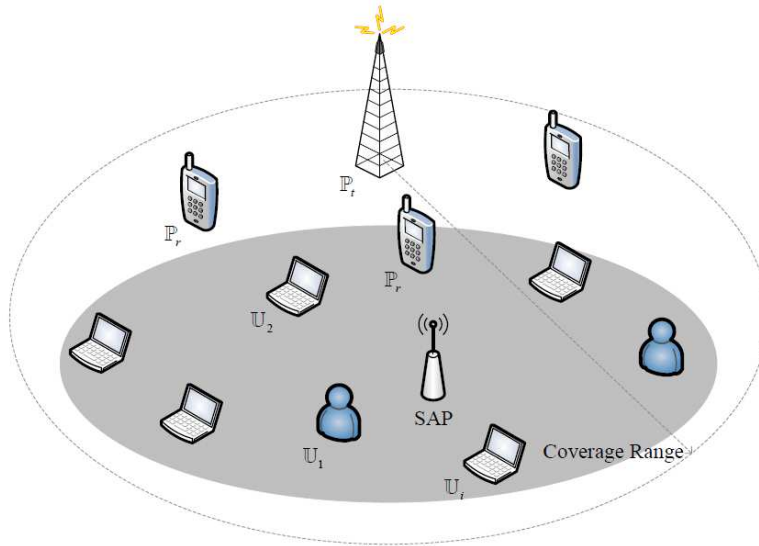


Figura 2.7: Modelo da rede CR de [CLMW11].

Quando um *PU* inicia a transmissão, os *SUs* devem evacuar o canal dentro de um tempo máximo de T_v segundos. Consequentemente, os *SUs* devem efetuar o *sensing* do meio com um período máximo de T_v . Assim, cada trama do *SU* com duração T_f ($T_f \leq T_v$)

deve conter uma fase de *sensing* T_s no início da trama e uma fase de transmissão T_d composta por M períodos de transmissão (TP). Cada TP inclui um tempo de transmissão T e um atraso de propagação T_p . O *sensing* efetuado pelos *SUs* da rede é baseado na energia do meio.

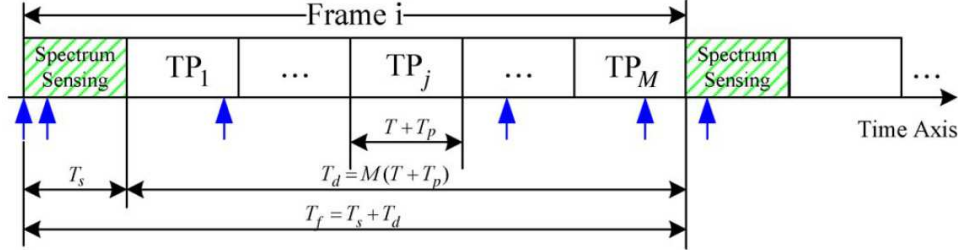


Figura 2.8: Estrutura da trama MAC do *Slotted CR-ALOHA*

Chen *et al.* [CLMW11] propõe um modelo de desempenho para o protocolo *Slotted CR-ALOHA*. Para a rede secundária, a carga gerada por cada *SU* é modelada através de uma distribuição de *Poisson* com uma taxa de geração média de λ_i pacotes por TP . Isto significa que o intervalo entre a geração de pacotes segue uma distribuição exponencial com média $\frac{1}{\lambda_i}$. Por outro lado, se se assumir que o tráfego gerado é igual para toda a rede secundária, então a carga total de pacotes (G) é de $N\lambda$.

Segundo os autores, os *PUs* e os *SUs* coexistem na mesma banda de espectro mas a rede primária não tem conhecimento da trama utilizada pelos *SUs*, pelo que existe a possibilidade de ocorrerem quatro hipóteses de comportamentos por parte do *PU* durante cada trama do secundário:

- \mathcal{H}_0 - o *PU* mantém-se inativo durante a fase de *sensing*;
- \mathcal{H}_1 - o *PU* mantém-se ativo durante a fase de *sensing*;
- \mathcal{H}_2 - o *PU* mantém-se inativo durante a fase de *sensing* mas acorda durante a fase de transmissão;
- \mathcal{H}_3 - o *PU* mantém-se inativo durante toda a trama de *SU*.

No caso de algum *SU* não detetar o sinal do *PU* sobre \mathcal{H}_1 ou transmitir sobre \mathcal{H}_2 , a rede secundária causa interferência sobre a primária. O parâmetro que mede este fenómeno chama-se fator de interferência. Por outro lado, é possível medir a capacidade de um *SU*

evacuar rapidamente o canal quando um *PU* acorda. Este parâmetro chama-se fator de agilidade. Este protocolo foi desenvolvido a partir do modelo *Slotted ALOHA* clássico e difere apenas no tempo de acesso discreto e na obrigação da proteção da rede primária. A estratégia do método *Slotted CR-ALOHA* é a seguinte:

- Se um *SU* detetar que um utilizador primário está inativo e se um pacote de dados chegar durante o *TP* número M da trama anterior ou durante a deteção de espectro da trama atual, este será enviado no início do primeiro *TP* desta trama. Se o pacote chegar durante o *TP* número j , sendo que $1 \leq j < M$, o pacote será transmitido no próximo *slot*.
- Se um *SU* detetar que existe um utilizador primário ativo, qualquer pacote que chegue dentro da trama atual será bloqueado até ao final desta trama e será, então, retransmitido após um tempo de contenção aleatório.
- Qualquer transmissão de dados só é enviada com sucesso quando apenas um *SU* transmite para um *slot TP*. Caso contrário, ocorre uma colisão e os pacotes nela envolvidos são retransmitidos, após um período aleatório, para evitar que volte a ocorrer outra colisão.
- Qualquer chegada de pacotes dentro do último *TP* (M) de uma trama será processada na próxima trama.

No artigo [CLMW11], é proposta uma expressão para o cálculo do débito (S) dependente do número de *SUs* (N) e do tempo de *sensing* (t). Os autores apresentam um modelo para P_{FA} e P_D onde assumem que P_{FA} decresce linearmente com t mas que P_D se mantém constante, o que se mostrou em [CLMW13] que não é totalmente válido. Nestas condições, concluem que, para $G \leq 1$ ($N = 25$ e $N = 50$), o sistema alcança um melhor desempenho quando t atinge o seu valor máximo ($t = T_S$). Todavia, o débito começa a diminuir com o aumento de *SUs* (para $N > 50$), provavelmente devido ao valor de janela de contenção usado. Para $G \leq 1$, valores reduzidos de P_{FA} aumentam as oportunidades de transmissão e permitem alcançar um melhor desempenho. Na realidade, ao aumentar o tempo de *sensing* reduz-se o tempo disponível para transmitir, mas a normalização realizada no artigo parece escamotear este efeito. Para $G > 1$, valores reduzidos de P_{FA}

agravam a carga do sistema, resultando em mais colisões e provocando uma degradação do desempenho do sistema. Por outro lado, quando $G > 1$ e $N > 50$, é possível manter o débito constante utilizando o valor de *sensing* ótimo t^* . Este trabalho mistura controlo de acesso com falso alarme; caso se tivesse estudado um parâmetro ajustável de acesso (janela de contenção *win*), não seria necessário analisar este efeito secundário na gestão da janela, que acaba por ter uma probabilidade de acesso efetiva de $(1 - P_{FA})/win$.

Em [CLMW11] mostra-se que o atraso médio de cada pacote (D) tende a crescer com o aumento da carga do sistema. Desta forma, D cresce linearmente para $G \leq 1$ e sofre um crescimento acelerado para $G > 1$, pois o sistema entra numa zona de saturação. No entanto, os autores conseguem atenuar o crescimento de D através da utilização de um valor ótimo t^* . Como t^* é menor do que T_S , provoca uma redução no evento de colisões, tal como acontece no débito. Este t^* representa o melhor equilíbrio possível entre o desempenho da operação de *sensing* e a disponibilização de banda para transmissão de dados.

Apesar do débito e do atraso médio dependerem do valor da janela de contenção, os autores não avaliam o modelo em função do seu valor, referem apenas que deve ser suficientemente grande para que não hajam colisões sucessivas. No entanto, este parâmetro é muito relevante, como se mostra nesta dissertação.

2.3.2 Protocolo *Cognitive Radio CSMA* (CR-CSMA)

O *Carrier Sense Multiple Access* (CSMA) 1-persistente é um protocolo de acesso aleatório ao meio que procura evitar colisões em redes de acesso múltiplo. Sempre que um nó pretende iniciar uma transmissão, verifica primeiro se o canal está isento de qualquer transmissão (incluindo transmissões de redes secundárias). Se o canal estiver ocupado, o nó monitoriza-o até que este se encontre livre e transmite imediatamente.

Em [CLMW09], é apresentado uma rede de *RC* baseada no protocolo *CSMA* 1-persistente. Mais uma vez, as redes primária e secundária partilham o mesmo canal de transmissão e a última não necessita de canal de controlo. O modelo do sistema é similar ao do protocolo Slotted CR-ALOHA (figura 2.7), onde existem apenas um *PU* e N *SUs* e é assumido que os *SUs* estão distribuídos dentro do alcance de transmissão do *PU* (figura

2.7). Apesar deste protocolo ser baseado em CSMA, os autores indicam que todos os *SUs* funcionam de um modo sincronizado entre eles, ou seja, este protocolo é considerado um protocolo baseado em *slots de tempo*.

Tal como no Slotted CR-ALOHA, assim que o *PU* começa a transmitir, os *SUs* devem evacuar o canal dentro de T_s segundos. Portanto, a rede secundária faz *sensing* ao meio com um período de pelo menos T_s , o que exige uma implementação em intervalos discretos. Além disso, cada trama T_f é constituída por uma fase de *sensing* de τ_{max} segundos e de transmissão de T_t segundos. A fase de transmissão é constituída por M períodos de transmissão (TPs) e cada TP inclui um tempo de transmissão de um pacote T e um tempo de propagação do pacote T_p . A estrutura da trama é idêntica à do *Slotted CR-ALOHA*, representada na figura 2.8.

O método de *sensing* utilizado é a deteção de energia no meio. Utiliza-se P_{noc} para denotar a probabilidade de um *SU* não ter colisões com o *PU* numa trama. P_{noc} deve ser mantido dentro de um valor de *threshold* \bar{P}_{noc} . Nesta condições, a probabilidade de um *SU* detetar a transmissão do *PU*, P_d , deve ser fixada em $\bar{P}_d = \bar{P}_{noc}^{\frac{1}{N}}$. Por outro lado, os autores também deduzem a expressão para a probabilidade de falso alarme e esta decresce continuamente em função do tempo de *sensing*, para valores de \bar{P}_d fixos.

Os autores propõem um modelo para o desempenho de *CR-CSMA* em [CLMW09], onde é assumido que a taxa de geração de pacotes de cada *SU* é dada por uma distribuição de *Poisson* independente com uma média de λ pacotes por TP . Desta forma, a taxa total de tráfego é dada por $G = N\lambda$. Cada TP é constituído por m *slots* de tempo (ST) e é assumido que cada *SU* transmite apenas no início de cada ST .

Posto isto, a estratégia da abordagem *CR-CSMA* é semelhante à abordagem tomada pelo *Slotted CR-ALOHA*, distinguindo-se apenas num pormenor importante: se um *SU* detetar que um utilizador primário está inativo e se o pacote chegar durante o TP número j , sendo que $1 \leq j < M$, acontecerá o seguinte:

- Se o canal estiver vago, o pacote será transmitido logo no próximo ST ;
- Se o canal estiver ocupado, o respetivo nó secundário monitoriza-o até que este esteja desocupado, transmitindo assim no início do primeiro ST livre.

Por fim, os autores fazem uma análise ao desempenho do modelo proposto através da

dedução da expressão do débito e do atraso médio de pacotes e constatam que os resultados obtidos melhoram ligeiramente o débito e o atraso médio de cada pacote face ao protocolo da secção 2.3.1. Esta melhoria de desempenho tem origem em dois motivos: i) neste protocolo, quando o *PU* não é detetado, os *SUs* esperam até ao próximo *ST* para tentar transmitir os pacotes recém-chegados, enquanto que no protocolo *Slotted CR-ALOHA*, os utilizadores secundários têm que aguardar pelo próximo *TP* para tentar aceder ao canal; ii) Adicionalmente, os *SUs* deste protocolo podem efetuar a deteção do meio durante a fase de transmissão, criando oportunidades de transmitir o pacote ainda no *TP* atual e não no próximo.

2.3.3 Protocolo *Slotted ALOHA P*-persistente

O protocolo *slotted Aloha P*-persistente estudado nesta dissertação foi analisado anteriormente em [HYS12]. Corresponde a um protocolo distribuído semelhante aos dois apresentados anteriormente, onde as redes primária e secundária partilham o mesmo canal único e os *SUs* encontram-se dentro do alcance do transmissor do *PU*. Em [HYS12] assume-se que existem N *SUs* e que todos eles têm a mesma taxa de chegada de pacotes, a mesma probabilidade P de aceder ao canal e modelam cada um com uma fila de espera do tipo $G/G/1/K^1$ com o mesmo tamanho, com uma filosofia *first-come first-served* (FCFS). A chegada de pacotes de cada *SU* é independente e modelada através de um processo de Bernoulli com uma taxa média de chegada de λ . Os pacotes podem ser rejeitados após a sua chegada quando a fila de espera de um *SU* está cheia. Este protocolo é baseado no método clássico do *Slotted ALOHA P*-persistente onde o tempo é dividido em *slots* e a transmissão de um pacote tem a duração de um *slot*, pelo que podem ser transmitidos vários pacotes durante o período de transmissão. Além disso, no caso de dois *SUs* acederem ao meio ao mesmo tempo, é considerado o efeito de captura no lado do recetor; ou seja, apenas a transmissão com mais potência ou a mais perto do recetor dos *SUs* é que é recebida caso o rácio entre a potência do sinal e a interferência seja superior a um limiar. Os outros pacotes cuja contenção falhou ou que não tiveram permissão para aceder ao meio permanecem na fila de espera do respetivo *SU* e esperam pela próxima trama para

¹ $G/G/1/K$ representa um sistema com uma única fila de serviço finita com dimensão k e uma filosofia de FCFS, com uma distribuição geral dos tempos de chegada e tempos de serviço.

serem transmitidos.

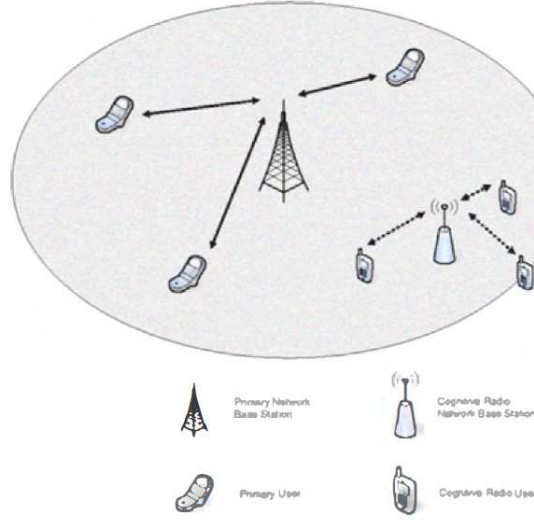


Figura 2.9: Topologia da rede do *Slotted CR-ALOHA* [HYS12].

Na figura 2.10, os SUs "ouvem" o canal da rede primária durante o seu período de *sensing*. De seguida, transmitem os seus pacotes durante um período de transmissão quando o canal está desocupado. Visto que o período de *sensing* e o período de transmissão são fixos neste modelo, os períodos de transmissão são tratados como *slots* de tempo para acesso ao canal. No início de cada trama, os *SUs* fazem *sensing* ao meio. Se o canal estiver ocupado, os *SUs* não têm permissão de aceder ao meio durante o próximo período de transmissão. Entretanto, todos os pacotes gerados são adicionados às listas dos respetivos utilizadores. Se o canal estiver livre, os *SUs* que possuírem pacotes na sua fila tentam aceder ao canal e transmitem com probabilidade de acesso P .

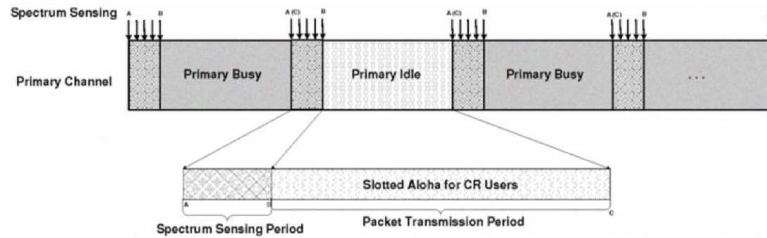


Figura 2.10: Slotted ALOHA para os *PUs* [HYS12].

Em [HYS12], os autores deduzem as expressões do débito e do atraso médio para cada pacote. No entanto, o modelo proposto não modela o tempo de atividade do primário mas apenas a probabilidade de uma trama estar ocupada pelo primário, considerando esta

probabilidade invariável no tempo e uniforme. No capítulo seguinte mostra-se que esta abordagem leva a erros significativos. Das expressões deduzidas, os autores observam que o débito aumenta e o atraso médio diminui à medida que a probabilidade do PU estar inativo (P_{idle}) aumenta. Além disso, se P_{idle} for constante e se variar a carga do sistema, os autores constataam que o desempenho do modelo também é diferente. Desta forma, o débito diminui e o atraso médio aumenta com uma carga total do sistema crescente devido ao facto de existirem mais pacotes a competir pelo acesso ao meio. Por outro lado, a probabilidade de acesso ao meio P também influencia o desempenho do sistema. No entanto, nos resultados apresentados os autores não mostram a sua influência.

2.3.4 Comparação entre os protocolos

Para terminar o capítulo 2, são comparados os protocolos *CR-Aloha* [CLMW11], *CR-CSMA* [CLMW09] e *Slotted Aloha P-persistente* [HYS12], apresentados anteriormente.

Os três protocolos analisados são distribuídos e descentralizados. Todos utilizam um *sensing* baseado deteção de energia do meio, pelo que necessitam que os utilizadores da rede secundária estejam sincronizados, caso contrário podem confundir outros SUs com o PU . Além disso, são todos baseados num único canal. Por outro lado, diferem na distribuição do *sensing*, que ocorre em cada *slot* para o *Slotted Aloha p-persistente*, e no início de uma trama com vários *slots*, nos dois casos restantes.

Os modelos propostos por [CLMW11] e [CLMW09] analisam principalmente a influência dos parâmetros de *sensing* no desempenho do sistema mas ignoram a influência da janela de contenção. Por outro lado, o último protocolo não considerou a duração do tempo de atividade de PU . Neste caso, quando o valor de P se aproxima de um, a hipótese da probabilidade do canal estar ocupado pelo PU ser assumida como constante e independente em dois *slots* consecutivos não é válida.

Na tabela 2.1 pode-se observar a síntese dos três protocolos.

Tabela 2.1: Comparação entre os três protocolos.

	CR-ALOHA	CR-CSMA	Slotted ALOHA P-persistente
Sensing baseado em detecção de energia	Sim	Sim	Sim
Canal único	Sim	Sim	Sim
Análise do parâmetro de sensing	Sim	Sim	Não
Análise da janela de contação	Não	Não	Sim
Ponderação da atividade do PU	Sim	Sim	Não

Capítulo 3

Descrição do modelo

3.1 Descrição do Sistema

Para a elaboração deste trabalho considerou-se um sistema de rádio cognitivo com apenas um *PU* e J *SUs* a aceder ao canal ($J \geq 1$) através de um protocolo *Slotted ALOHA P-persistente*. O *PU* pode aceder ao meio sempre que pretender e, se o canal estiver ocupado por *SUs*, estes devem desimpedi-lo. Os utilizadores secundários são *SUs* equipados apenas com um *transceiver*, pelo que não podem fazer o *sensing* do canal ao mesmo tempo que transmitem informação, sendo obrigados a efetuar a operação de *sensing* periodicamente. Como os *SUs* não conseguem distinguir as transmissões dos *PU*s das suas, é de grande importância que os *SUs* estejam sincronizados, fazendo a operação de *sensing* durante o mesmo intervalo de tempo, para que só existam *PU*s a aceder ao canal. Caso contrário, seria possível o aparecimento de eventos de falso alarme no sistema devido a deteções de sinais de *SUs* confundidos com o sinal do *PU*. De qualquer modo, como os *SUs* baseiam as suas operações de *sensing* em deteção de energia, podem ser originadas falsas deteções do sinal do *PU* devido ao ruído com origem térmica ou noutros sinais presentes no meio. Os *SUs* acedem ao canal com probabilidade P quando detetam que o *PU* está inativo durante a fase inicial de *sensing* de um *slot*. No entanto, é possível que existam colisões quando dois ou mais *SUs* decidem iniciar uma transmissão ao mesmo tempo. Também é possível existirem colisões entre os *SUs* e o *PU* se algum *SU* não conseguir detetar a transmissão do *PU* e decidir transmitir. No caso de ocorrer uma colisão, os pacotes têm que ser novamente transmitidos até que sejam recebidos com sucesso.

Cada trama de SU é composta por um período de *sensing* com duração de T_S^{SU} slots seguido de um período de transmissão com duração de T_D^{SU} slots, tendo a duração total de T_F^{SU} slots. A duração de cada *slot* é dada pelo período de amostragem do canal.

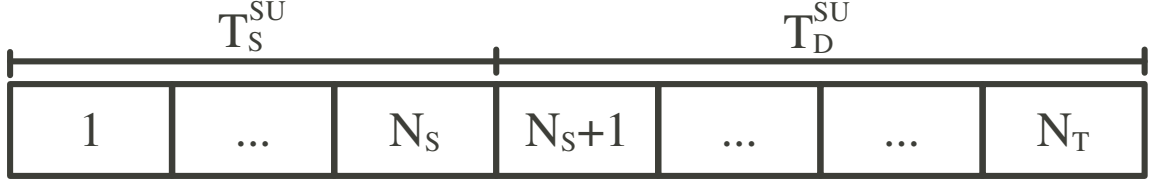


Figura 3.1: Trama de um SU .

Os primeiros N_S slots definem a duração do período de *sensing* e os restantes slots assinalam a duração do período de transmissão (de $N_S + 1$ até N_T).

Por outro lado, os tempos de atividade do PU estão divididos entre período ativo T_F^{PU} e período inativo $\overline{T_F^{PU}}$. O acesso dos SUs deve minimizar a perturbação que introduz nos PU s. De forma a garantir que os SUs detetam rapidamente a atividade dos PU s, cada ciclo dos utilizadores secundários deve ser significativamente mais curto do que o período ativo/inativo do PU ($T_F^{SU} < \min(T_F^{PU}, \overline{T_F^{PU}})$) [GS07]. Desta maneira, o tamanho dos ciclos dos SUs deve ser escolhido de acordo com a seguinte expressão: $T_F^{SU} = \frac{\min(T_F^{PU}, \overline{T_F^{PU}})}{\alpha}$ onde $\alpha > 1$ [LFO⁺13]. De salientar que, para a elaboração deste trabalho, foi utilizado um valor de $\alpha = 15$ para limitar a interferência entre o PU e os SUs [LFO⁺13].

A taxa de chegada de pacotes dos SUs é modelada através de um processo de *Poisson* com um ritmo médio de λ pacotes de dados por trama. Os pacotes são guardados numa fila de espera até iniciarem a transmissão. A transmissão de um pacote pode demorar um tempo variável que depende do meio estar a ser usado pelo PU , do valor da variável P que condiciona o acesso, e de possíveis colisões. Nesta dissertação, o sistema é modelado por uma fila de espera do tipo M/G/1 [Kle76][Dai05].

3.1.1 Comportamento do utilizador primário

Neste trabalho, foi considerado que o PU tem um comportamento não constante, isto é, o PU pode aceder ou desocupar o canal aleatoriamente durante uma trama dos SUs . Posto isto, assume-se que todos os pacotes transmitidos durante um *slot* onde ocorre uma alteração de estado de PU não são recebidos com sucesso.

A modelação da atividade do *PU* foi gerada a partir da cadeia de *Markov* ilustrada abaixo, constituída por dois estados e respetivas probabilidades de transição (probabilidade de alteração de estado ativo para inativo (π_{10}) ou vice-versa (π_{01}) e probabilidade de se manter no mesmo estado (π_{00} ou π_{11})). Para $T_F^{PU} < \overline{T_F^{PU}}$ as probabilidades são dadas por [LFO⁺13]:

$$\begin{aligned}\pi_{10} &= \frac{1}{\alpha N_T} \\ \pi_{11} &= 1 - \pi_{10} \\ \pi_{01} &= \frac{P_\beta}{\alpha N_T(1-P_\beta)} \\ \pi_{00} &= 1 - \pi_{01}\end{aligned}\tag{3.1}$$

onde P_β é a probabilidades do *PU* se manter ativo.

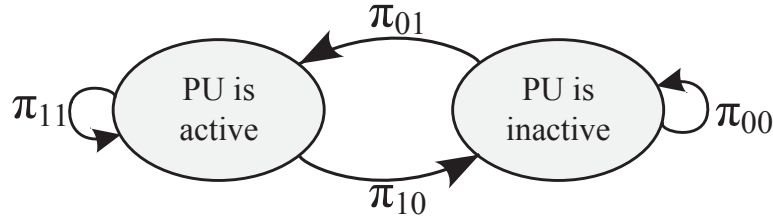


Figura 3.2: Cadeia de Markov do comportamento do *PU*.

Para calcular a probabilidade de falso alarme (P_{FA}) e de deteção (P_D) é necessário calcular a probabilidade da energia detetada por um *SU* ser superior ao valor de limiar. Quando se dá a alteração de estado do *PU* durante a operação de *sensing* dos *SUs*, podem ocorrer dois cenários distintos:

- **Alteração do estado ativo para inativo de *PU*** - Durante os primeiros *slots* do período de *sensing* o *PU* está ativo (até N_G *slots*). Nos restantes *slots*, o estado do *PU* fica inativo.
- **Alteração do estado inativo para ativo de *PU*** - Ao contrário do cenário anterior, o *PU* altera o seu estado de inativo para ativo após N_G *slots* pelo que de $N_G + 1$ até N_S , o *PU* mantém-se ativo.

Assim, destes cenários surgem duas novas hipóteses, H_{01} e H_{10} , para juntar às já existentes H_0 e H_1 abordadas no capítulo 2.

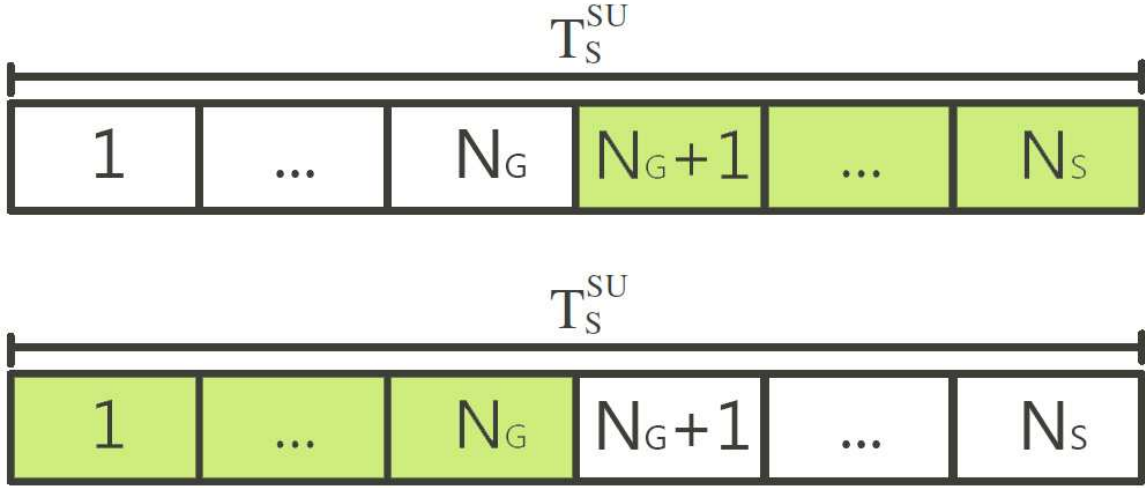


Figura 3.3: Dois cenários possíveis na alteração de estados do PU .

O sinal detetado H_{01} é representado pela seguinte expressão:

$$Y_{H_{01}} = \sum_{k=1}^{N_G} |w(k)|^2 + \sum_{k=N_G+1}^{N_S} |w(k) + s(k)|^2 \quad (3.2)$$

onde o lado esquerdo da soma é o sinal detetado durante o início do período de *sensing* até ao *slot* N_G enquanto PU está ativo e, por isso, é equivalente à expressão da hipótese H_0 . Como esta parte da soma segue uma distribuição central chi-quadrado com N_G graus de liberdade e assumindo que N_G é suficientemente grande, o Teorema do Limite Central é aplicável e esta distribuição chi-quadrado é aproximadamente igual à distribuição normal $Y_{H_{01}}^a \sim \mathcal{N}(N_G, 2N_G)$. O lado direito da soma, que é o sinal detetado durante os *slots* finais do período de *sensing* onde o PU já está inativo, é igual à expressão de H_1 . Esta segue uma distribuição chi-quadrado não-centralizada com $N_S - N_G$ graus de liberdade e um parâmetro de não-centralidade κ e, tal como a anterior, através do Teorema do Limite Central, é aproximadamente igual a uma distribuição normal $Y_{H_{01}}^b \sim \mathcal{N}(N_S - N_G + \kappa, 2(N_S - N_G + 2\kappa))$. Como é assumido que $Y_{H_{01}}^a$ e $Y_{H_{01}}^b$ são independentes e identicamente distribuídas, obtêm-se $Y_{H_{01}} \sim \mathcal{N}(N_S + \kappa, 2N_S + 4\kappa)$. Por fim, como PU apenas está ativo durante $N_S - N_G$ *slots*, κ é dado por $\kappa = (N_S - N_G)\kappa'$ [LFO⁺13].

Seguindo o mesmo raciocínio para a hipótese H_{10} , obtemos as expressões para as

quatro hipóteses possíveis da deteção de sinal:

$$Y = \begin{cases} \mathcal{N}(N_S, 2N_S), & H_0 \\ \mathcal{N}(N_S + \kappa' N_G, 2(N_S + 2\kappa' N_G)), & H_{10} \\ \mathcal{N}(N_S + \kappa'(N_S - N_G), 2N_S + 4\kappa'(N_S - N_G)), & H_{01} \\ \mathcal{N}(N_S + \kappa, 2(N_S + 2\kappa)), & H_1 \end{cases} \quad (3.3)$$

A partir de 3.3 é possível calcular as probabilidades de falso alarme e de deteção. Assume-se que só ocorre uma deteção falsa quando se verificam as hipóteses H_0 e H_{10} porque o PU não está ativo no fim do período de *sensing*. Posto isto, com as probabilidades condicionadas $P_{FA}^{H_{10}}$ e $P_{FA}^{H_0}$, deduzidas em [LFO⁺13], e com as probabilidades de ocorrerem eventos H_0 e H_{10} , $\mathbb{P}\{H_{10}\}$ e $\mathbb{P}\{H_0\}$, deduz-se a probabilidade de falso alarme P_{FA} :

$$P_{FA} = P_{FA}^{H_{10}} \mathbb{P}\{H_{10}\} + P_{FA}^{H_0} \mathbb{P}\{H_0\}, \quad (3.4)$$

onde

$$P_{FA}^{H_{10}} = Pr(Y > \gamma | H_{10}) = \mathcal{Q}\left(\frac{\gamma - N_S - N_G \kappa'}{\sqrt{4N_G \kappa' + 2N_S}}\right), \quad (3.5)$$

$$P_{FA}^{H_0} = Pr(Y > \gamma | H_0) = \mathcal{Q}\left(\frac{\gamma - N_S}{\sqrt{2N_S}}\right) \quad (3.6)$$

e γ é o valor de limiar de decisão que já foi introduzido anteriormente. No capítulo 4 é apresentado o método de cálculo de γ .

Através dos mesmos métodos e assumindo que só ocorre uma deteção do sinal de PU quando se verificam as hipóteses H_1 e H_{01} , também se calculou a probabilidade de deteção dada por:

$$P_D = P_{FA}^{H_{01}} \mathbb{P}\{H_{01}\} + P_{FA}^{H_1} \mathbb{P}\{H_1\}, \quad (3.7)$$

onde

$$P_D^{H_{01}} = Pr(Y > \gamma | H_{01}) = \mathcal{Q}\left(\frac{\gamma - N_S - \kappa'(N_S - N_G)}{\sqrt{2N_S + 4\kappa'(N_S - N_G)}}\right), \quad (3.8)$$

$$P_D^{H_1} = Pr(Y > \gamma | H_1) = \mathcal{Q}\left(\frac{\gamma - (N_S + \kappa)}{\sqrt{2(N_S + 2\kappa)}}\right) \quad (3.9)$$

Na simulação são considerados os quatro estados (H_0 , H_{10} , H_{01} e H_1). No entanto, como é assumido que os pacotes transmitidos durante um *slot* onde ocorre uma alteração

de estado de PU não são recebidos com sucesso, os valores relativos a H_{01} e H_{10} não são considerados para fins de cálculo de P_{FA} e P_D no modelo analítico. Assim, considera-se que a probabilidade de sucesso e, conseqüentemente, o débito são nulos em tais casos.

3.2 Modelação do sistema

O sistema proposto é um processo estocástico que tem uma complexidade de modelação significativa, resultante de ser o resultado do comportamento aleatório de vários elementos aleatórios independentes. Desta forma, foram adotadas aproximações e simplificações para se conseguir modelar o seu comportamento.

A modelação do sistema foi realizada em três abordagens sucessivas, de complexidade crescente. Para cada abordagem são analisadas as falhas e imprecisões e é identificada a hipótese usada que mais contribuiu para os erros obtidos. A secção encontra-se dividida em três partes, cada uma representando uma abordagem analisada para a formulação das expressões do modelo.

3.2.1 Primeira abordagem - Modelação *slot a slot*

A primeira abordagem modelou o sistema considerando a dinâmica das filas de espera dos *SUs slot a slot*, semelhante à adotada em [HYS12]. Como hipóteses simplificadoras, consideraram-se um sistema com carga uniforme (i.e. todos os *SUs* geram carga com a mesma distribuição) modelada por um processo de *Poisson*, os *slots* estão ocupados pelo *PU* com uma probabilidade $P\{\tau = 1\}$ e estão livres com uma probabilidade $P\{\tau = 0\}$.

Modelo

Modela-se o número de pacotes da fila de espera que um *SU* contém no *slot* m , representado por q_m , através de uma cadeia de *Markov* com 3 estados. Os estados desta cadeia representam os três acontecimentos possíveis durante um *slot* de *SU*:

- Se a fila de espera estiver vazia ($q_m = 0$), os novos pacotes são-lhe adicionados;
- Se a fila de espera não estiver vazia ($q_m > 0$), o *PU* estiver inativo ($\tau = 0$), não houver falso alarme ($1 - P_{FA}$), o *SU* decidir transmitir (P) e se não houver erros

de transmissão $(1 - P_{err})$, então há um pacote que é transmitido com sucesso e os novos pacotes são adicionados à fila de espera;

- Se a fila de espera não estiver vazia ($q_m > 0$), o PU estiver ativo ($1 - (\tau = 0)$) ou houver falso alarme (P_{FA}) ou o SU não decidir transmitir ($1 - P$) ou se houver erros de transmissão (P_{err}), a informação não é transmitida com sucesso, sendo apenas acrescentados à fila os novos pacotes.

A cadeia de *Markov* pode ser expressa por:

$$q_{m+1} = \begin{cases} v_0 & \text{if } q_m = 0 \\ q_m + v_0 & \text{if } q_m > 0, \tau = 1, (1 - (1 - P_{FA}) \times P \times (1 - P_{err})) \\ q_m + v_0 - 1 & \text{if } q_m > 0, \tau = 0, (1 - P_{FA}) \times P \times (1 - P_{err}) \end{cases} \quad (3.10)$$

onde v_0 é a variável aleatória que representa o número de pacotes de dados recebidos durante um *slot*, q_m representa o número de pacotes na fila de espera, τ representa o estado do PU , P_{FA} representa a probabilidade de falso alarme, P representa a probabilidade de um utilizador decidir transmitir e P_{err} é a probabilidade de existirem colisões na transmissão do pacote, e calcula-se através da equação $P_{err} = 1 - (1 - P(1 - P_{QE})(1 - P_{FA}))^{J-1}$. Admite-se que não há erros no canal, i.e. que todos os pacotes transmitidos isoladamente são recebidos com sucesso.

A cadeia de *Markov* tem um papel fundamental para a modelação da dinâmica do sistema. A partir do estado estacionário obtido com o modelo podem-se calcular as expressões da probabilidade da fila de espera dos SUs estar vazia, representada por P_{QE} , do tamanho médio da fila de espera e do atraso médio na fila. No entanto, é importante salientar que o modelo exibido considera que a probabilidade da fila de espera estar vazia é invariante no tempo (independentemente do número de *slot*, P_{QE} é constante). Na realidade, a fila de espera vai crescendo e diminuindo ao longo do tempo. Sempre que o PU acede, os SUs não podem aceder ao canal e, à medida que vão chegando pacotes de dados para serem transmitidos, estes vão sendo adicionados à fila de espera, que cresce continuamente. Por outro lado, quando o PU altera o seu estado para inativo, os SUs começam transmitir os dados e as suas filas de espera começam a esvaziar. Desta forma, não se define o valor esperado de q_m de uma forma estrita, pois o processo q_m não é

ergódico na média. Mas é possível utilizar a abordagem proposta por Cesàro [SW94], que corresponde a definir o valor esperado como o limite do valor médio de q_m num intervalo quando a duração do intervalo tende para infinito, conforme está representado na equação 3.11.

$$Q(z) = \lim_{m \rightarrow +\infty} E[z^{q_m+1}]. \quad (3.11)$$

Deste modo, introduz-se um erro na modelação do sistema, exemplificado com a evolução da fila de espera ao longo do tempo, representada na figura 3.4: aproxima-se o valor de P_{QE} pelo seu valor médio no tempo. Esta aproximação funciona pior quando maior for a oscilação da probabilidade ao longo do intervalo. Neste caso, funciona bem quando o sistema tem uma carga baixa ($P_{QE} \approx 1$) ou quando está saturado ($P_{QE} = 0$), mas apresenta um erro maior para cargas intermédias.

Esta aproximação foi usada por outros autores [Bia00] para modelar sistemas com o protocolo *Distributed Coordinated Function* do IEEE 802.11, que também apresentam esta variação temporal durante o período de *backoff*. A validade da utilização do valor médio de acesso (i.e. $1 - P_{QE}$) foi comprovada em [HDML08] para 802.11, que mostrou que apenas revelava um desvio significativo quando o tamanho da fila de espera crescia para sistemas não saturados. No entanto, [HDML08] também mostra que a hipótese das filas de espera dos vários nós serem independentes e identicamente distribuídas não era totalmente válida, embora o resultado final da sua utilização desse bons resultados (e.g. [OK09] com 802.11 com tráfego *unicast* e *broadcast*).

O primeiro passo para calcular a probabilidade da fila de espera estar vazia P_{QE} , é aplicar a função geradora de probabilidades [BGdMT06] à cadeia de *Markov* (equação 3.11):

$$\begin{aligned} Q(z) = & E[z^{v_0}] \times P\{q_m = 0\} + \\ & + E[z^{v_0}] \times z^{-1} \times E[z^{q_m | q_m > 0}] \times P\{q_m > 0\} \times P\{\tau = 0\} \times (1 - P_{FA}) \times P \times (1 - P_{err}) + \\ & + E[z^{v_0}] \times E[z^{q_m | q_m > 0}] \times P\{q_m > 0\} \times (1 - (1 - P_{FA} \times P\{\tau = 0\}) \times P \times (1 - P_{err}), \end{aligned} \quad (3.12)$$

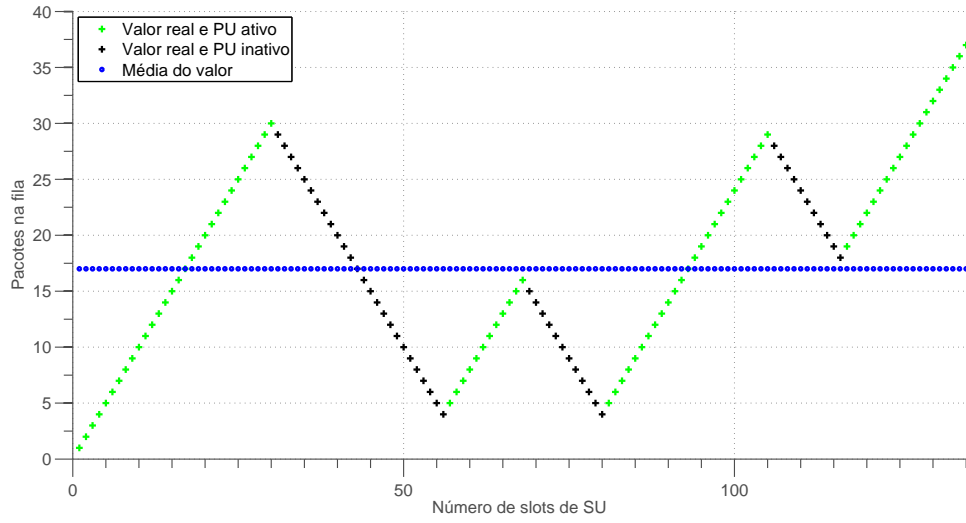


Figura 3.4: Valor instantâneo e valor médio do número de pacotes na fila de espera ao longo do tempo.

Visto que v_0 é um processo de *Poisson*, o valor esperado de z^{v_0} é igual a:

$$E[z^{v_0}] = \sum_{k=1}^{\infty} \frac{e^{-\lambda} \times \lambda^k}{k!} z = e^{\lambda(z-1)} = V_0(z), \quad (3.13)$$

e

$$E[z^{q_m | q_m > 0}] = \sum_{i=1}^{\infty} z^i P\{q_m = i | q_m > 0\} = \frac{Q(z) - P_{QE}}{1 - P_{QE}}. \quad (3.14)$$

Substituindo as expressões acima na equação 3.12 e resolvendo em ordem a $Q(z)$ obtemos a seguinte expressão:

$$Q(z) = \frac{V(z)P_{QE}(1 - z^{-1})(1 - \mathbb{P}\{\tau = 1\})(1 - P_{FA})P(1 - P(1 - P_{QE}))^{J-1}}{1 - V(z)(1 - (1 - z^{-1})(1 - \mathbb{P}\{\tau = 1\})(1 - P_{FA})P(1 - P(1 - P_{QE}))^{J-1})}. \quad (3.15)$$

De seguida, para se obter P_{QE} , utilizou-se a normalização da função densidade [BGdMT06] que é dada por $G(1) = 1$. Para resolver a indeterminação $G(1) = \frac{0}{0}$ aplicou-se a regra de *Cauchy* (3.16) e por fim, através de métodos numéricos, obteve-se os valores de P_{QE} :

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \quad (3.16)$$

$$\lambda = (1 - P_{QE})(1 - \mathbb{P}\{\tau = 1\})(1 - P_{FA})P(1 - P(1 - P_{QE}))^{J-1} \quad (3.17)$$

$$G_{tot_{sec}} = J \times P(1 - P_{QE}) \times (1 - P(1 - P_{QE}))^{J-1} \times ((1 - \mathbb{P}\{\tau = 1\}) \times (1 - P_{FA})). \quad (3.18)$$

Por fim, obteve-se a expressão do débito de um SU (3.18) onde:

- A primeira parte da expressão $(P(1 - P_{QE})(1 - P_{FA}))$ representa a probabilidade do SU transmitir;
- A segunda parte da expressão $((1 - P(1 - P_{FA})(1 - P_{QE}))^{J-1})$ consiste na probabilidade de mais nenhum SU transmitir;
- A terceira parte da expressão $(1 - \mathbb{P}\{\tau = 1\})$ compreende a probabilidade do PU não aceder ao canal.

Avaliação do desempenho

De forma a validar o modelo, foi desenvolvido um simulador para o sistema, descrito no capítulo 4.

$P_{\{\tau=1\}}$	N_S	N_T	γ	J	P	P_D	P_{FA}
0.502565	42	425	77.817634	2	0.5	0.997650	0.006177

Tabela 3.1: Parâmetros de configuração da 1ª abordagem.

A configuração referida na tabela 3.1 foi utilizada tanto no simulador como no modelo. As cargas reais geradas e usadas na simulação também foram utilizadas como parâmetro do modelo.

A figura 3.5 representa a probabilidade da fila de espera estar vazia P_{QE} quando o PU está inativo estimada com o modelo e medida pelo simulador. É possível verificar que o modelo apresenta um valor mais elevado do que os resultados medidos no simulador. Este desvio deve-se ao facto de se considerar, nesta abordagem, um comportamento de independência *slot* a *slot* (i.e. PU pode não estar ativo num *slot*, alterar o seu estado para ativo no *slot* seguinte e, por último, voltar à inatividade no *slot* a seguir). Na verdade, existe uma forte correlação nos *slots* consecutivos em que o PU se mantém ativo.

A figura 3.6 representa o débito de um SU obtido com o simulador e com o modelo. A figura mostra que o modelo acompanha o resultado das simulações à medida que a taxa de chegada de pacotes aumenta, mas devido ao modelo sobrestimar a probabilidade P_{QE} , observa-se que o débito estimado pelo modelo atinge um valor máximo maior que o medido na simulação.

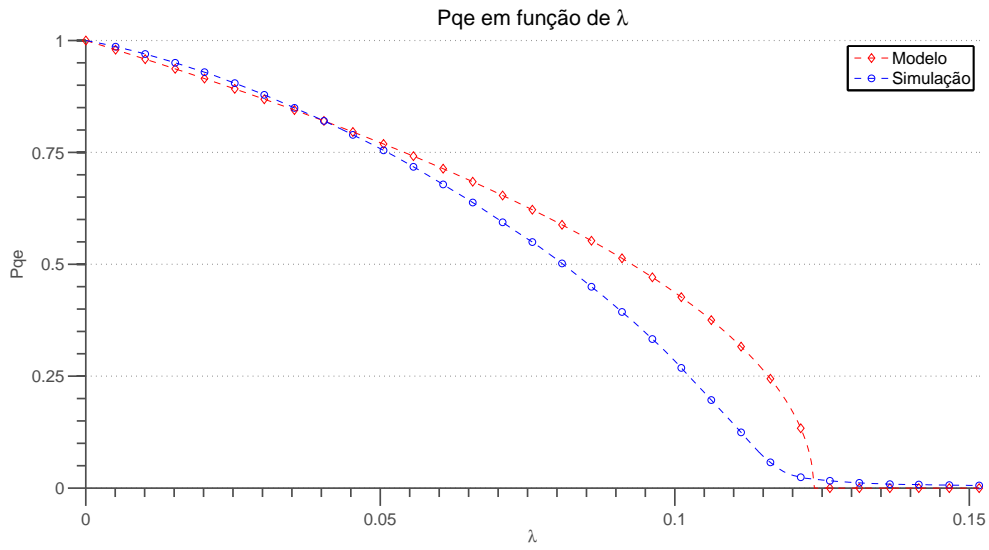


Figura 3.5: Resultados da probabilidade da fila de espera estar vazia obtidos com o modelo e com o simulador.

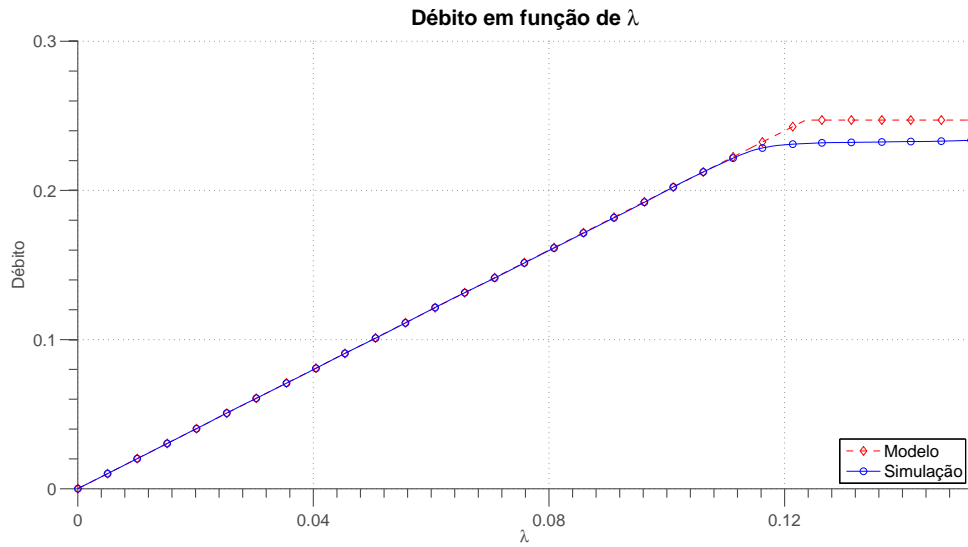


Figura 3.6: Débito por SU obtido com o modelo e com o simulador.

Assim, para corrigir estes desvios, concluiu-se que é necessário modelar corretamente a correlação entre *slots* que existe na probabilidade do PU estar ativo, $\pi_{1,1}$, ou inativo, $\pi_{0,0}$, visto que estes dois parâmetros influenciam significativamente todo o desempenho do sistema.

3.2.2 Segunda abordagem - Modelação de tempo de PU

A principal conclusão do primeiro modelo proposto foi que é necessário modelar a correlação entre *slots* do estado do *PU*, o que pode ser realizado considerando o estado do *PU* na cadeia de *Markov*. Para além disso, como a evolução da fila de espera com o *PU* ativo só depende da duração do período ativo (nunca há pacotes recebidos com sucesso), é possível agrupar o conjunto de *slots* com o *PU* ativo num único estado da cadeia de *Markov*. Desta forma, tal como anteriormente, a segunda abordagem modela o sistema considerando a dinâmica da fila de espera dos *SUs* *slot* a *slot* com o *PU* inativo, mas usa outro estado para quando o *PU* está ativo. De resto, manteve-se a hipótese de existir uma carga uniforme na rede.

Modelo para P_{QE}

Na cadeia de *Markov* mantiveram-se os três estados da abordagem anterior para modelar a dinâmica da independência dos *slots* quando $PU = 0$, enquanto se criou um novo estado quando $PU = 1$, representado na figura 3.7, onde se agrupam todos os *slots* dos *SUs* enquanto o *PU* está ativo. A carga média acumulada durante esta sequência de *slots* é modelada através da variável aleatória v_1 .

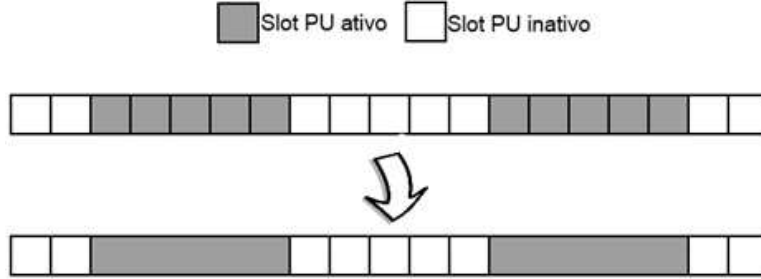


Figura 3.7: Agregação de *slots* com *PU* ativo na segunda abordagem.

Esta sequência de *slots* é vista apenas como uma transição. Desta forma, a cadeia foi atualizada para:

$$q_{m+1} = \begin{cases} v_0 & \text{if } PU = 0, q_m = 0 \\ q_m + v_0 & \text{if } PU = 0, (1 - (1 - P_{FA}) \times P \times !P_{err}) \\ q_m + v_0 - 1 & \text{if } PU = 0, (1 - P_{FA}) \times P \times !P_{err} \\ q_m + v_1 & \text{if } PU = 1 \end{cases}, \quad (3.19)$$

onde $P\{PU = 1\}$ mede a probabilidade de o estado de $PU = 1$ (estar ativo) ser atingido na cadeia de *Markov*. Desta forma, define-se $P\{PU = 1\} = \frac{1}{1+E[\tau_0]}$ e $P\{PU = 0\} = 1 - P\{PU = 1\}$.

No final de um *slot* de *SU* com o *PU* inativo:

- Se a fila de espera estiver vazia ($q_m = 0$), os novos pacotes aparecidos durante o *slot* são adicionados a esta;
- Se a fila de espera não estiver vazia ($q_m > 0$), se houver falso alarme (P_{FA}) ou se o *SU* decidir não transmitir ($1 - P$) ou se houver erros de transmissão, não há pacotes transmitidos com sucesso e os novos pacotes aparecidos durante o *slot* são adicionados à fila de espera;
- Se a fila de espera não estiver vazia ($q_m > 0$), não houver falso alarme ($1 - P_{FA}$), o *SU* decidir transmitir (P) e se não houver erros de transmissão, então é transmitido um pacote com sucesso e são acrescentados os novos pacotes.
- Contrariamente aos casos anteriores, se o *PU* estiver ativo ($PU = 1$), os pacotes recebidos durante todo o período ativo do *PU* são adicionados à fila de espera.

Tal como anteriormente, desenvolveram-se as expressões da probabilidade da fila de espera estar vazia quando *PU* está inativo, P_{QE} , através da função geradora de probabilidades.

$$Q(z) = E[z^{q_m+1}], \quad (3.20)$$

obtendo-se:

$$\begin{aligned} Q(z) = & E[z^{q_m+v_1}] \times P\{PU = 1\} + \\ & + E[z^{q_m+v_0}] \times P\{PU = 0\} \times P\{q_m > 0\} \times (1 - (1 - P_{FA} \times P \times (1 - P_{err}) + \\ & + E[z^{q_m+v_0-1}] \times P\{PU = 0\} \times P\{q_m > 0\} \times (1 - P_{FA} \times P \times (1 - P_{err}) + \\ & + E[z^{v_0}] \times P\{PU = 0\} \times P\{q_m = 0\}. \end{aligned} \quad (3.21)$$

Tal como foi realizado na primeira abordagem, substituíram-se as expressões e probabilidades pelas respetivas fórmulas (equações 3.13 e 3.14). A função geradora da variável

aleatória v_1 contabiliza a carga gerada pelo processo de *Poisson* com média λ , e é calculada com:

$$E[z^{v_1}] = \sum_{l=1}^{\infty} (1 - \pi_{1,1}) \pi_{1,1}^{l-1} \sum_{k=0}^{\infty} \frac{e^{-\lambda(l+1)} (\lambda(l+1))^k}{k!} z^k = \frac{(1 - \pi_{1,1}) V_0(z)^2}{1 - \pi_{1,1} V_0(z)} = V_1(z) \quad (3.22)$$

onde $\pi_{1,1}^1$ é a probabilidade do próximo *slot* de *SU* ser um *slot* com o *PU* ativo. Resolvendo em ordem a $Q(z)$, obtém-se a seguinte expressão:

$$Q(z) = \frac{(1 - PU) V_0(z) P_{QE} (1 - z^{-1}) P_{OK}}{1 - V_1(z) PU - (1 - PU) V_0(z) (1 - (1 - z^{-1}) P_{OK})} \quad (3.23)$$

onde P_{OK} , que corresponde à probabilidade de transmitir um pacote com sucesso, resulta da conjunção da probabilidade de não haver falso alarme, do *SU* transmitir e de os restantes $J - 1$ *SUs* não transmitirem e é dado por:

$$P_{OK} = P(1 - P_{FA})(1 - P(1 - P_{FA})(1 - P_{QE}))^{J-1}. \quad (3.24)$$

Utilizando novamente a normalização da função densidade obteve-se uma indeterminação ($G(1) = \frac{0}{0}$). Mais uma vez, aplicou-se a regra de Cauchy e, através de métodos numéricos, obteve-se a expressão 3.25.

$$(1 - P_{QE}) P_{OK} = \frac{\lambda}{1 - PU} \left(1 + \frac{PU}{1 - \pi_{1,1}}\right), \quad (3.25)$$

A duração média dos períodos onde o *PU* está ativo, $E[\tau_1]$, é:

$$E[\tau_1] = \sum_{l=0}^{\infty} (l + 2) \pi_{1,1}^l (1 - \pi_{1,1}) = \frac{2 - \pi_{1,1}}{1 - \pi_{1,1}}, \quad (3.26)$$

onde l é o número de *slots* com atividade do *PU* e $\pi_{1,1}$ é a probabilidade do próximo *slot* ser um *slot* com o *PU* ativo. Como no último *slot* de inatividade antes da sequência de *slots* com o *PU* ativo e o primeiro a seguir à sequência existe uma forte probabilidade de haver falha na deteção de atividade do *PU* (isto é, nos *slots* com transição de estado existe uma forte probabilidade de não existirem transmissões de sucesso), estes *slots* são

¹A definição de $\pi_{1,1}$ é aprofundada, com o auxílio da dedução da duração média dos períodos de atividade do *PU*, ainda nesta secção.

contabilizados como *slots* de atividade do *PU* (como pode ser observado na figura 3.8). Desta forma, na equação 3.26 são acrescentados à expressão estes dois *slots* extra ($l + 2$).

Por outro lado, com a perda destes dois *slots* de inatividade, a expressão da duração média dos períodos onde o *PU* está inativo é dada por:

$$E[\tau_0] = \sum_{l=0}^{\infty} l \pi_{0,0}^l (1 - \pi_{0,0}) = \frac{\pi_{0,0}}{1 - \pi_{0,0}} \quad (3.27)$$

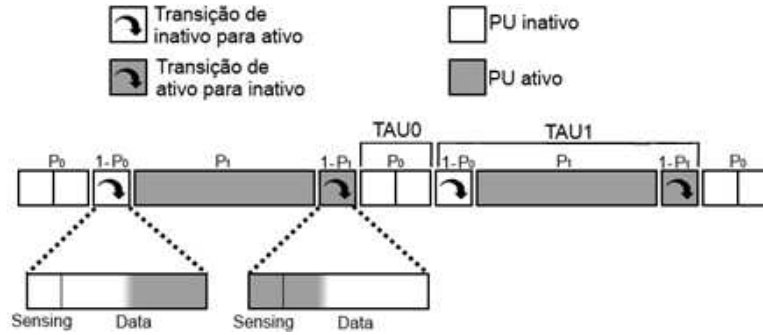


Figura 3.8: Slots com PU ativo e inativo. Representação de τ_0 e τ_1 .

Modelo para débito

A modelação do débito total para J *SUs* do sistema foi construída através da multiplicação da probabilidade de um pacote ser recebido com sucesso (P_{OK}) com as probabilidades do *PU* estar inativo ($1 - P_U$) e da fila de espera de um determinado *SU* não estar vazia ($1 - P_{QE}$).

$$S = J(1 - P_{QE})P_{OK}(1 - P_U). \quad (3.28)$$

Modelo para atraso

O tempo médio de atraso de um pacote enviado pelos *SUs* pode ser decomposto em tempo de serviço e tempo de espera na fila. Através da fórmula de *Pollaczek-Khinchine* (equação 3.30) [BG92], o tamanho médio de uma fila de espera do tipo M/G/1 relaciona-se com a distribuição do tempo de serviço e tempo de espera. De seguida, utiliza-se o teorema de *Little* [LG08] para obter o atraso médio de um pacote W . Este teorema é dado

por:

$$W = \frac{L_{SU}}{\lambda}, \quad (3.29)$$

onde L_{SU} é o tamanho médio da fila de espera e pode ser calculado através da fórmula de *Pollaczek-Khinchine*:

$$L_{SU} = \rho + \frac{\rho^2 + \lambda^2 \text{Var}(S)}{2(1 - \rho)}. \quad (3.30)$$

λ representa a taxa de chegada de um processo de *Poisson*, S é a distribuição do tempo de serviço, o ρ é a taxa de utilização da rede (é igual à taxa de chegada λ multiplicada pelo tempo médio de serviço ($E[S]$)) e $\text{Var}(S)$ é a variância do tempo de serviço. Assim, ρ é dado por:

$$\rho = \lambda \times E[S]. \quad (3.31)$$

Sabe-se que, para um pacote ser recebido com sucesso, um SU pode ter que transmiti-lo l vezes, até que seja recebido corretamente. Posto isto, $E[S]$ é dado por:

$$E[S] = \sum_{l=1}^{\infty} E[S_1 | \text{SU fez } l \text{ transmissões}] (1 - P_{Fail}) P_{Fail}^{l-1}. \quad (3.32)$$

Como as variáveis aleatórias da equação 3.32 são independentes, logo:

$$\begin{aligned} E[S] &= \sum_{l=0}^{\infty} l E[S_1] (1 - P_{Fail}) P_{Fail}^{l-1} = \\ &= E[S_1] (1 - P_{Fail}) \sum_{l=1}^{\infty} l P_{Fail}^{l-1} \end{aligned} \quad (3.33)$$

onde S_1 é uma variável aleatória que define a duração de uma retransmissão e $E[S_1]$ é dado por:

$$\begin{aligned} E[S_1] &= E[S_1 | \pi_{1,1}] \mathbb{P}_{\{PU=1\}} + E[S_1 | 1 - \pi_{1,1}] \mathbb{P}_{\{PU=0\}} = \\ &= E[S_1 | \pi_{1,1}] PU + (1 - PU). \end{aligned} \quad (3.34)$$

$E[S_1|\pi_{1,1}]$, que corresponde ao número médio de retransmissões sucessivas até um pacote ser enviado com sucesso, é dado por:

$$\begin{aligned} E[S_1|\pi_{1,1}] &= \sum_{m=0}^{\infty} mE[\tau_1](1 - \pi_{0,0})^{m-1}\pi_{0,0} = \\ &= \frac{E[\tau_1]}{\pi_{0,0}}. \end{aligned} \quad (3.35)$$

Por fim, substituindo as equações 3.35 e 3.36 em 3.34, obtém-se:

$$E[S] = \frac{P_{Fail}}{1 - P_{Fail}} \left(\frac{P_U}{\pi_{0,0}} E[\tau_1] + 1 - P_U \right) \quad (3.36)$$

e P_{Fail} é igual a $1 - P_{OK}$.

Por outro lado, para calcular $Var(S)$ utilizou-se a probabilidade de se ter l retransmissões. A variância, por definição, é dada por $E[(S - E[S])^2]$. Para simplificar o modelo, admitiu-se a hipótese de que a variância de cada retransmissão é independente, a variância da soma de variáveis aleatórias independentes é a soma das suas variâncias. Utilizou-se esta propriedade para deduzir as equações 3.37 e 3.39. Na realidade, a variância não é totalmente independente. Esta simplificação é uma das fontes de erros que afeta a estimação do atraso médio de pacotes no modelo e que se reflete num valor inferior ao real. No entanto, para evitá-lo, seria necessário calcular a covariância das várias combinações possíveis.

$$\begin{aligned} Var(S) = E[(S - E[S])^2] &= \sum_{l=1}^{\infty} lE[(S_1 - E[S_1])^2](1 - P_{Fail})P_{Fail}^{l-1} = \\ &= \frac{P_{Fail}}{1 - P_{Fail}} E[(S_1 - E[S_1])^2]. \end{aligned} \quad (3.37)$$

Assim, a variância do tempo de serviço $Var(S)$ é dada pelo valor esperado da variância da duração da transmissão de um pacote até ter sucesso $Var(S_1)$.

$$\begin{aligned} Var(S_1) = E[(S_1 - E[S_1])^2] &= E[(S_1 - E[S_1])^2|\pi_{1,1}]\mathbb{P}\{PU = 1\} + \\ &+ E[(S_1 - E[S_1])^2|1 - \pi_{1,1}]\mathbb{P}\{PU = 0\} = \\ &= E[(S_1 - E[S_1])^2|\pi_{1,1}]PU. \end{aligned} \quad (3.38)$$

A variância da duração de uma transmissão sabendo que o PU está ativo no próximo *slot* $E[(S_1 - E[S_1])^2 | \pi_{1,1}]$ foi deduzida através do valor esperado das variâncias dos períodos sucessivos onde o PU está ativo e é dada por:

$$E[(S_1 - E[S_1])^2 | \pi_{1,1}] = \sum_{m=1}^{\infty} m E[(\tau_1 - E[\tau_1])^2] (1 - \pi_{0,0})^{m-1} \pi_{0,0} = \frac{E[(\tau_1 - E[\tau_1])^2]}{\pi_{0,0}}, \quad (3.39)$$

A variância de um único período onde o PU está ativo $Var(\tau_1)$ é dada por:

$$\begin{aligned} Var(\tau_1) &= E[(\tau_1 - E[\tau_1])^2] = \sum_{l=0}^{\infty} (l + 2 - E[\tau_1])^2 \pi_{1,1}^l (1 - \pi_{1,1}) = \\ &= (1 - \pi_{1,1}) \left(\sum_{l=0}^{\infty} (l + 2)^2 \pi_{1,1}^l - 2E[\tau_1] \sum_{l=0}^{\infty} (l + 2) \pi_{1,1}^l + E[\tau_1]^2 \sum_{l=0}^{\infty} \pi_{1,1}^l \right) = \\ &= \frac{(1 + \pi_{1,1})\pi_{1,1}}{(1 - \pi_{1,1})^2} + \frac{4\pi_{1,1}}{1 - \pi_{1,1}} + 4 - \frac{2E[\tau_1]\pi_{1,1}}{1 - \pi_{1,1}} - 4E[\tau_1] + E[\tau_1]^2 = \\ &= \frac{(1 + \pi_{1,1})\pi_{1,1}}{(1 - \pi_{1,1})^2} + 2(2 - E[\tau_1]) \frac{\pi_{1,1}}{(1 - \pi_{1,1})} + (2 - E[\tau_1])^2. \end{aligned} \quad (3.40)$$

Por fim, simplificando e substituindo as expressões em $Var(S)$, obtém-se:

$$Var(S) = \frac{P_{Fail}}{1 - P_{Fail}} \times P_U \times \frac{\frac{1 + \pi_{1,1}}{(1 - \pi_{1,1})^3} + 2(2 - E[\tau_1]) \frac{\pi_{1,1}}{(1 - \pi_{1,1})^2} + (2 - E[\tau_1])^2}{\pi_{0,0}} \quad (3.41)$$

Validação

Para validar o modelo comparou-se o desempenho medido no simulador com o previsto pelo modelo. Os parâmetros utilizados no simulador e no modelo para obter os gráficos desta secção encontram-se expostos na tabela 3.2.

PU	N_S	N_T	γ	J	P	$\pi_{1,1}$	$\pi_{0,0}$	P_D	P_{FA}
0.502565	42	425	77.817634	2	0.5	0.935762	0.935209	0.997642	0.006173

Tabela 3.2: Parâmetros de configuração da 2ª abordagem.

A figura 3.9 ilustra a variação do débito com a carga, mostrando que os valores do modelo e do simulador são quase coincidentes.

Os resultados da probabilidade da fila de espera de um SU estar vazia, P_{QE} , representados na figura 3.10 mostram um bom ajuste entre o modelo e as simulações do sistema,

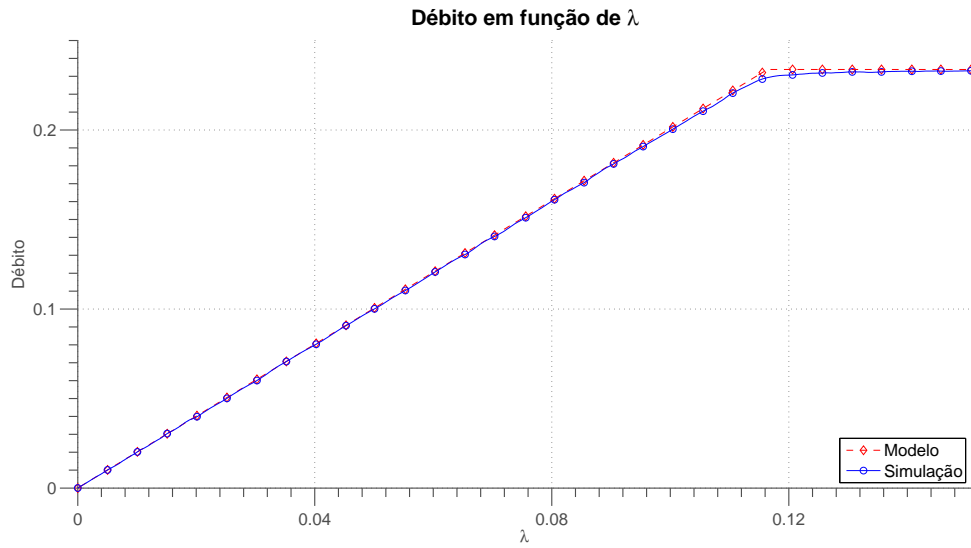


Figura 3.9: Resultados do débito obtidos com o modelo e com o simulador.

exceto no limiar da saturação, onde $P_{QE} = 0$: o modelo mostra uma queda abrupta para $\lambda = 0.115$ pacotes/slot mas na simulação a saturação apenas ocorre para $\lambda \approx 0.130$ pacotes/slot. Este erro resulta das aproximações usadas, nomeadamente de se ter considerado que a probabilidade de erro numa retransmissão de um pacote era igual à probabilidade de erro da primeira transmissão. A influência deste último erro fica mais clara quando se visualiza o erro no atraso.

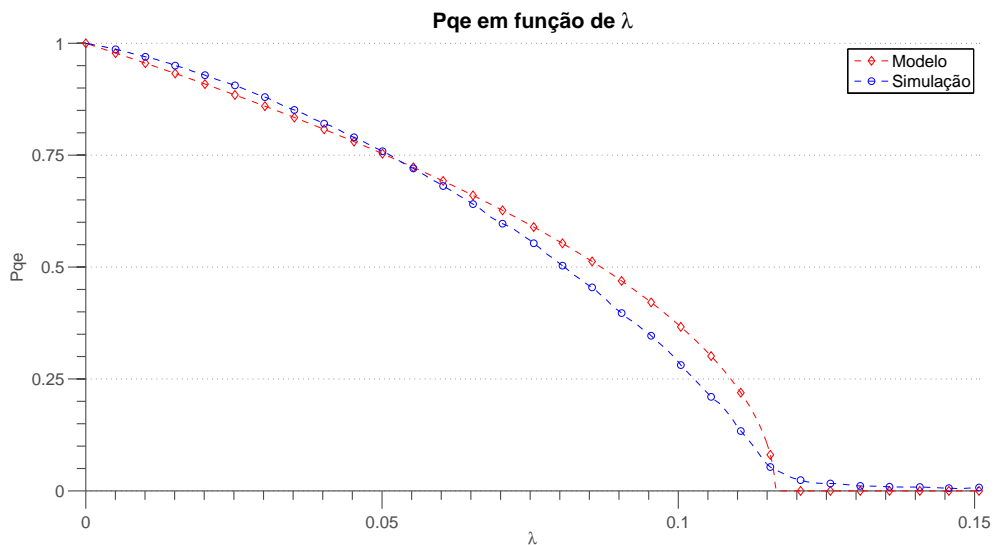


Figura 3.10: Resultados da probabilidade da fila de espera estar vazia obtidos com o modelo e com o simulador.

A figura 3.11 representa os valores de atraso, e mostra que o modelo apresenta um

erro significativo. O modelo falha ao não capturar o crescimento acelerado do atraso à medida que a carga aumenta, que resulta do aumento do número de colisões (e de erros de transmissão) à medida que o número de nós com pacotes na fila aumenta. Além disso, também existe um valor de *offset* nos resultados da simulação que não está a ser modelado.

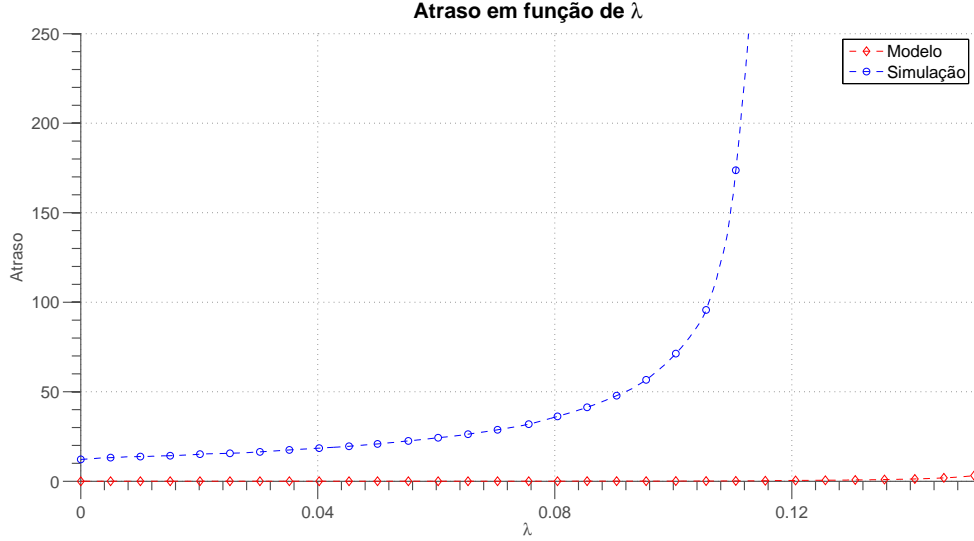


Figura 3.11: Resultados do atraso médio por pacote obtidos com o modelo e com o simulador.

Apesar de o modelo ter apresentado um erro relativamente baixo para P_{QE} e praticamente nulo para o débito no cenário representado anteriormente, com $P = 0.5$, na figura 3.12 mostra-se que para valores de P mais elevados, as estimativas do modelo são cada vez mais imprecisas. Assim, após uma breve reflexão, verificou-se que não foi ponderada corretamente no modelo a hipótese de um pacote falhar sucessivamente. Ou seja, identificou-se que a probabilidade de erro não é independente do número de transmissão, mas existe uma correlação relevante de P_{Err} com ser ou não a primeira retransmissão.

3.2.3 Terceira abordagem - modelação de tempo de PU e de erros

O modelo de canal considerado nesta abordagem é, como anteriormente, um modelo com efeito de captura de canal, onde um erro ocorre apenas quando mais do que um SU ou PU transmitem para o canal. Quando ocorre um erro numa transmissão para o canal com N SUs a transmitirem, vai existir uma probabilidade de o erro persistir no próximo *slot* sem transmissões do PU com um valor maior ou igual a $1 - (1 - P(1 - P_{FA}))^N$.

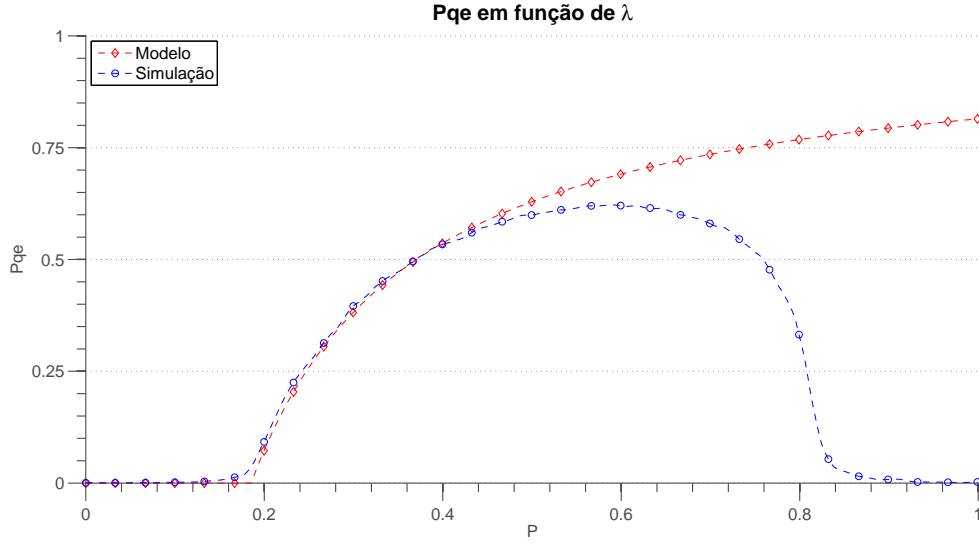


Figura 3.12: Resultados da probabilidade da fila de espera estar vazia obtidos do modelo e do simulador em função de P .

Quando o valor de P é elevado, a probabilidade do erro persistir no tempo é também elevada. Quanto maior for o número de retransmissões, maior é o tempo decorrido e maior é a probabilidade de chegarem mais pacotes à fila dos SUs , crescendo o número de SUs a concorrerem no acesso ao meio. No entanto, este crescimento da probabilidade de erro com N é lento, e a sua modelação obrigaria a introduzir estados na cadeia de *Markov* para modelar o número médio de pacotes após um, dois, etc., erros sucessivos. De forma a limitar a complexidade do modelo proposto, decidiu-se desprezar esta variação e considerar a aproximação que a probabilidade de erro numa retransmissão de um pacote é constante e diferente da probabilidade de erro na primeira transmissão de um pacote.

Modelo para P_{QE}

A probabilidade de erro para um pacote transmitido pela primeira vez é dada pela equação abaixo, que é equivalente às expressões de erro das abordagens anteriores.

$$P_{err_1} = 1 - (1 - P(1 - P_{FA})(1 - P_{QE}))^{J-1}. \quad (3.42)$$

Deste modo, a probabilidade de sucesso para uma primeira transmissão de um pacote

depende da probabilidade de não haverem erros de transmissão ($1 - P_{err}$) e é dada por:

$$\varsigma_1 = P(1 - P_{FA})(1 - P_{err_1}). \quad (3.43)$$

No entanto, quando a primeira transmissão de um pacote falha, há pelo menos um SU com pacotes na sua fila de espera. Como se assume que as filas de espera dos SUs são independentes e identicamente distribuídas, a probabilidade de erro depende da distribuição dos SUs com pacotes na fila de espera. Por isso, a função massa de probabilidade condicionada do número de SUs com a fila com pacotes, Φ_k , é dada por:

$$\begin{aligned} \Phi_k &= \mathbb{P}\{q = k | P_{err_1}\} = \frac{\mathbb{P}\{q = k\}}{\mathbb{P}\{q > 1\}} \\ &= \binom{J-1}{k} \left(\frac{1 - P_{QE}}{P_{QE}} \right)^k \frac{(P_{QE})^{J-1}}{1 - (P_{QE})^{J-1}} \end{aligned} \quad (3.44)$$

A probabilidade de erro para uma transmissão posterior à primeira é calculada com base na probabilidade de não haver outros SUs a transmitir considerando todas as possíveis combinações do número de SUs com pacotes para enviar. Esta probabilidade de erro é dada por:

$$\begin{aligned} P_{err_n} &= \sum_{k=1}^{J-1} \Phi_k (1 - (1 - P(1 - P_{FA}))^k) \times (1 - P(1 - P_{FA})(1 - P_{QE}))^{J-1-k} = \\ &= 1 - \frac{P_{QE}^{J-1} (1 - P(1 - P_{FA})(1 - P_{QE}))^{J-1}}{1 - P_{QE}^{J-1}} \times \left(\left(1 + \left(\frac{1 - P_{QE}}{P_{QE}} \right) \left(\frac{1 - P(1 - P_{FA})}{1 - P(1 - P_{FA})(1 - P_{QE})} \right) \right)^{J-1} - 1 \right) \end{aligned} \quad (3.45)$$

Deste modo, a probabilidade de sucesso para as sucessivas retransmissões de pacotes (ς_n) é dada pela equação (3.46) e depende de não haverem erros de transmissão posteriores aos primeiros ($1 - P_{err_n}$).

$$\varsigma_n = P(1 - P_{FA})(1 - P_{err_n}). \quad (3.46)$$

A partir das probabilidades de sucesso nas transmissões e de erro sucessivo, é possível deduzir as expressões do tempo de serviço que serão necessárias para, mais uma vez, obter a probabilidade da fila de espera estar vazia P_{QE} . A duração esperada de um período com PU ativo, $E[\tau_1]$, é calculada utilizando (3.26), contabilizando-se os *slots* de transição entre a atividade e inatividade como *slots* não utilizáveis pelos SUs . De forma semelhante, a duração esperada do período com PU inativo, $E[\tau_0]$, é calculado utilizando (3.27). Para facilitar a leitura do documento, as duas

equações são reproduzidas de seguida.

$$\begin{aligned} E[\tau_1] &= \sum_{l=0}^{\infty} (l+2)(\pi_{1,1})^l(1-\pi_{1,1}) = \frac{2-\pi_{1,1}}{1-\pi_{1,1}}, \\ E[\tau_0] &= \sum_{l=0}^{\infty} l(\pi_{0,0})^l(1-\pi_{0,0}) = \frac{\pi_{0,0}}{1-\pi_{0,0}}. \end{aligned}$$

Em média, um SU espera um número de *slots* igual a $E[\Delta_{SU}^1]$ até à chegada de um *slot* em que o SU pode transmitir com sucesso. Caso a duração do período ativo do PU seja da ordem de grandeza da duração de um *slot* de SU , este tempo pode incluir várias transições do PU entre ativo e inativo, sendo calculado usando,

$$\mathbb{E}[\Delta_{SU_1}] = \pi_{0,0} \sum_{l=0}^{\infty} (l+1)(1-\pi_{0,0})^l(1+l\mathbb{E}[\tau_1]) = \frac{2-\pi_{0,0}-\pi_{1,1}}{(1-\pi_{1,1})\pi_{0,0}} = 1 + \frac{\mathbb{E}[\tau_1]}{\mathbb{E}[\tau_0]} \quad (3.47)$$

O tempo de serviço esperado de um pacote, $E[\Delta_{SU}]$, é calculado contabilizando o número médio de retransmissões necessárias até que o pacote seja recebido com sucesso.

$$\begin{aligned} E[\Delta_{SU}] &= \mathbb{E}[\Delta_{SU_1}]_{\varsigma_1} + (1-\varsigma_1) \left(\sum_{k=0}^{\infty} (k+1) \mathbb{E}[\Delta_{SU_1}]_{\varsigma_n} (1-\varsigma_n)^{k-1} \right) \\ &= \frac{1+\varsigma_n-\varsigma_1}{\varsigma_n} \mathbb{E}[\Delta_{SU_1}] \end{aligned} \quad (3.48)$$

Ao contrário das abordagens anteriores, a probabilidade da fila de espera de um SU estar vazia é determinada independentemente do PU estar ativo ou inativo. Para o cálculo da probabilidade da fila de espera estar vazia, P_{QE} , é necessário obter o rácio de utilização da rede (3.31), onde λ é a carga média por SU . Assim, para uma fila M/G/1 com um tempo de serviço invariante no tempo, [Tak62] mostra que $P_{QE} = 1 - \lambda E[\Delta_{SU}]$ para $\lambda E[\Delta_{SU}] < 1$ e que $P_{QE} = 0$ para $\lambda E[\Delta_{SU}] \geq 1$. Assim, utilizando um método numérico é possível calcular o valor de P_{QE} resolvendo a equação (3.49). Importa realçar que este valor é uma aproximação, uma vez que tanto P_{QE} como $E[\Delta_{SU}]$ variam no tempo, definindo-se a probabilidade e o valor esperado apenas com a definição de *Cèsaro*. O erro cometido é maior para níveis de carga intermédios, onde a oscilação da ocupação da fila de espera é maior entre as fases com PU ativo e com PU inativo.

$$1 - P_{QE} = \lambda E[\Delta_{SU}] \quad (3.49)$$

Modelo para débito

O débito agregado dos J SUs do sistema pode ser calculado considerando a fração dos *slots* em que se transmite efetivamente a informação com sucesso, descontando sucessivamente os *slots*

usados pelo PU , sem pacotes na fila ou desperdiçados com retransmissões. O débito agregado é dado por:

$$S_a = J(1 - P_{QE})(1 - P_U) \frac{E[\Delta_{SU}^1]}{E[\Delta_{SU}]}, \quad (3.50)$$

onde P_U é a probabilidade do PU estar ativo num *slot*, dada por:

$$P_U = \frac{\mathbb{E}[\tau_1]}{\mathbb{E}[\tau_1] + \mathbb{E}[\tau_0]} = \frac{(2 - \pi_{1,1})(1 - \pi_{0,0})}{2 - \pi_{1,1} - \pi_{0,0}}. \quad (3.51)$$

Desenvolvendo a expressão anterior, obtém-se $S_a = J\lambda$ desde que o sistema não esteja saturado. Além do débito agregado dos SUs , é possível calcular o débito agregado máximo do sistema, que ocorre quando o sistema está saturado, para $P_{QE} = 0$. Assim:

$$S_{sat} = J(1 - P_U)P(1 - P_{FA})(1 - P(1 - P_{FA}))^{J-1}. \quad (3.52)$$

Modelo para atraso

O tempo médio de atraso para um pacote enviado pelos SUs inclui o tempo de serviço, $\mathbb{E}[\Delta_{SU}]$ e o tempo de espera na fila, cuja soma pode ser estimada através da fórmula de *Pollaczek-Khinchine* [BG92], dado assumir-se que a fila de espera é do tipo M/G/1. Na formulação apresentada anteriormente para o modelo do sistema, assumiu-se que os pacotes chegam sempre no fim de um *slot* de SU , quando o PU está inativo, ou no fim do último *slot* em que o PU está ativo. Na realidade, o tempo de chegada dos pacotes à fila de espera é uma variável aleatória contínua, sendo necessário acrescentar ao atraso total o tempo médio desde que o pacote chega à fila até ao alinhamento no tempo com o modelo, designado de tempo de inatividade (*Vacation time*). Assumindo que a duração máxima esperada do tempo de inatividade é $E[V]$ e que a distribuição de chegada de pacotes é uniforme no tempo, o tempo médio de inatividade é dado por $E[V]/2$.

Desta forma, o tempo de atraso médio do sistema para um pacote de SU é:

$$E[\Delta_T] = \mathbb{E}[\Delta_{SU}] + \frac{\lambda \mathbb{E}[(\Delta_{SU})^2]}{2(1 - \lambda \mathbb{E}[\Delta_{SU}])} + \frac{\mathbb{E}[V]}{2}. \quad (3.53)$$

O segundo momento do atraso de serviço $\mathbb{E}[(\Delta_{SU})^2]$ é necessário para calcular o tempo de espera na fila e é determinado através da variância do tempo de serviço $\sigma_{\Delta_{SU}}^2$ do SU . Por outro lado, esta variância depende do número de *slots* irrelevantes (ocupados pelo PU), que por sua vez depende da variância dos *slots* irrelevantes $\sigma_{\tau_1}^2$. Estas expressões são dadas por:

$$\mathbb{E}[(\Delta_{SU})^2] = (\mathbb{E}[\Delta_{SU}])^2 + \sigma_{\Delta_{SU}}^2, \quad (3.54)$$

$$\sigma_{\Delta_{SU}}^2 = \sigma_{\Delta_{SU}^1}^2 \varsigma_1 + (1 - \varsigma_1) \left(\sum_{k=0}^{\infty} (k+1) \sigma_{\Delta_{SU}^1}^2 \varsigma_n (1 - \varsigma_n)^{k-1} \right) = \frac{1 + \varsigma_n - \varsigma_1}{\varsigma_n} \sigma_{\Delta_{SU}^1}^2, \quad (3.55)$$

$$\sigma_{\Delta_{SU}^1}^2 = \pi_{0,0} \sum_{l=0}^{\infty} l(1 - \pi_{0,0})^l \sigma_{\tau_1}^2 = \frac{1 - \pi_{0,0}}{\pi_{0,0}} \sigma_{\tau_1}^2, \quad (3.56)$$

$$\begin{aligned} \sigma_{\tau_1}^2 &= \mathbb{E}[(\tau_1)^2] - (\mathbb{E}[\tau_1])^2 = \sum_{l=0}^{\infty} (l+2)^2 (\pi_{1,1})^l (1 - \pi_{1,1}) - (\mathbb{E}[\tau_1])^2 \\ &= \frac{4 - 3\pi_{1,1} + (\pi_{1,1})^2}{(1 - \pi_{1,1})^2} - \frac{1}{(1 - \pi_{1,1})^2} = \frac{3 - 3\pi_{1,1} + (\pi_{1,1})^2}{(1 - \pi_{1,1})^2}. \end{aligned} \quad (3.57)$$

O tempo de inatividade é obtido através do tempo médio de espera até ao fim de uma trama de SU ou até ao fim de uma sequência de atividade do PU , dado por

$$\mathbb{E}[V] = P_U \mathbb{E}[\tau_1] + (1 - P_U) = 1 + \frac{(2 - \pi_{1,1})(1 - \pi_{0,0})}{(1 - \pi_{1,1})(2 - \pi_{1,1} - \pi_{0,0})}. \quad (3.58)$$

Validação

A validação do modelo foi realizada considerando valores de referência obtidos com o simulador. Foram considerados os parâmetros apresentados na tabela 3.3.

PU	N_S	N_T	γ	J	P	$\pi_{1,1}$	$\pi_{0,0}$	P_D	P_{FA}
0.105644	42	425	77.812909	5	0.2	0.934448	0.992234	0.998343	0.001083

Tabela 3.3: Parâmetros de configuração da 3ª abordagem.

A figura 3.13 representa a variação do valor de P_{QE} em função da carga. Observa-se um desvio nos valores do modelo que é maior para valores de carga intermédia, como era esperado. O modelo usa a aproximação de considerar um valor "médio" para P_{QE} . Estes resultados mostram que o valor não corresponde à média temporal medida no simulador. Mas os dois valores coincidem para cargas muito baixas ou na saturação, onde a variação de P_{QE} ao longo do tempo é minimizada.

A figura 3.14 ilustra a variação do débito com a carga, mostrando que os valores do modelo e do simulador são coincidentes. Este resultado é previsível, atendendo à abordagem usada no desenho do modelo. A figura 3.15 representa o atraso total por pacote, mostrando que os valores estimados apresentam apenas uma ligeira imprecisão, demonstrando a sua validade. Quando o sistema atinge a saturação, as filas de espera dos SUs tendem a crescer rapidamente com o tempo total de simulação - enquanto o modelo assume um valor de infinito.

Claramente, observa-se que este terceiro modelo permite obter estimativas válidas para o tempo de atraso e para o débito, embora apresente um desvio visível no valor de P_{QE} .

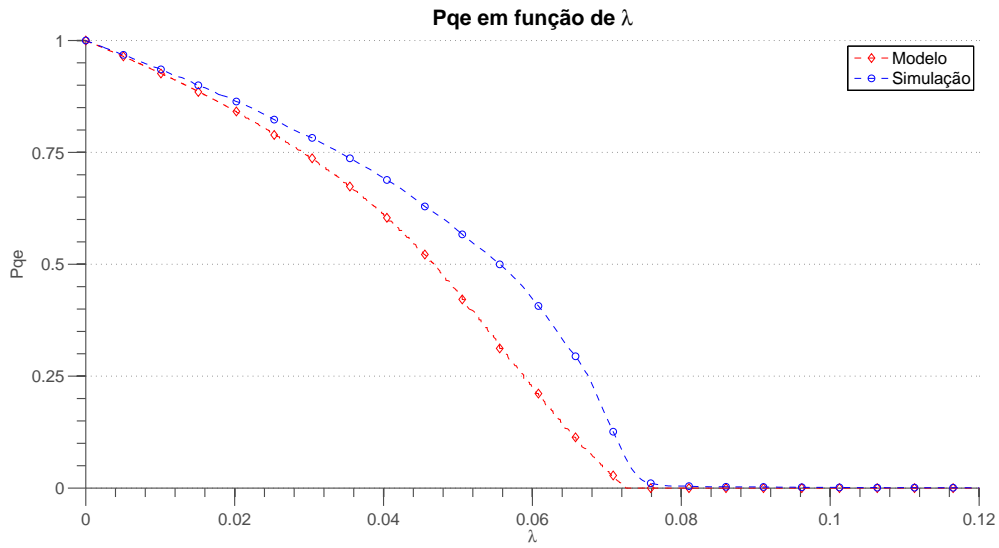


Figura 3.13: P_{QE} em função da carga para o modelo e o simulador.

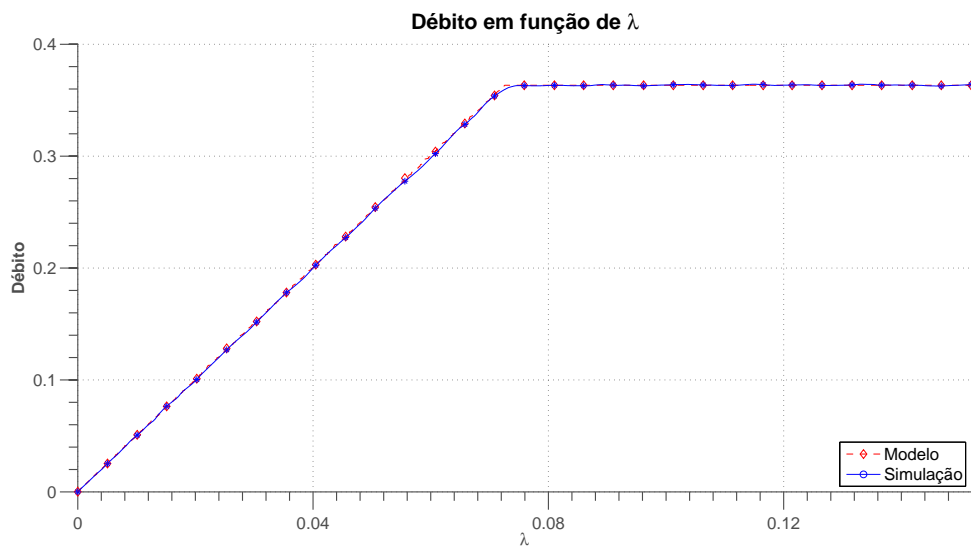


Figura 3.14: Resultados do débito obtidos do modelo e do simulador.

Síntese do percurso das abordagens

A tabela 3.4 contém o sumário das abordagens consideradas e das limitações encontradas nas três abordagens estudadas. É relevante observar que a abordagem seguida em [HYS12] usa a primeira abordagem estudada, sofrendo dos desvios apresentados.

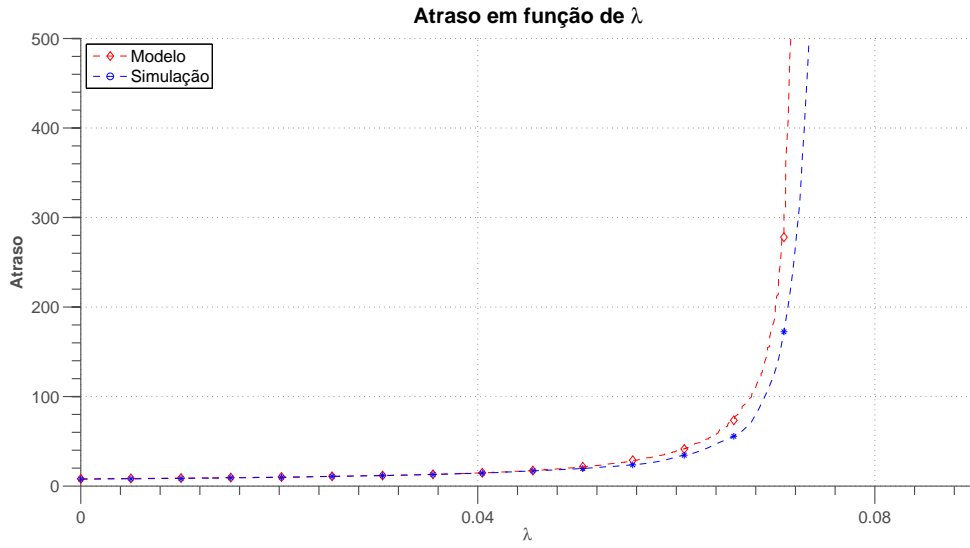


Figura 3.15: Resultados do atraso médio de um pacote obtidos do modelo e do simulador.

Abordagem	Princípio da modelação	Notas
1ª	Slot a slot de SU	Desvio do P_{QE} , débito e atraso
2ª	Correlação entre slots de SU	Imprecisão na modelação do atraso
3ª	Correlação entre ser ou não a 1ª transmissão	Pequeno desvio do P_{QE} . Precisão dos valores de débito/atraso modelados Trade-off entre P_{QE} e débito/atraso.

Tabela 3.4: Síntese das abordagens.

3.3 Otimização do Modelo

Ao longo das abordagens, verificou-se que o valor de P influencia os resultados do débito e do atraso. Esta secção analisa o desempenho global para vários valores de P e apura o valor de P que maximiza o débito do sistema e que minimiza o atraso médio de pacotes.

Quando o valor de P cresce de 0 até 1, o débito aumenta até atingir um valor máximo, pois reduz-se o número de *slots* de contenção (sem acesso) até o *SU* transmitir. Quando o número de *SUs* é maior que 1, com o crescimento de P a partir do máximo, o número de *slots* desperdiçados devido a colisões começa a subir provocando a redução de débito, que atinge um valor nulo com $P = 1$. Como o declive do débito, $\frac{\partial S_{sat}}{\partial p}$, é nulo quando o valor máximo é atingido, é possível definir o valor de P que maximiza o débito do sistema, P_{sat}^* , através da seguinte expressão:

$$\frac{\partial S_{sat}}{\partial p} = \left(1 - \varrho - p \frac{\partial \varrho}{\partial p}\right) (1 - Jp(1 - P_{FA})(1 - \varrho)) = 0, \quad (3.59)$$

Resolvendo a equação 3.59 obtém-se a expressão 3.60, que pode ser resolvida através de métodos numéricos.

$$p_{sat}^*(1 - \varrho) = \frac{1}{J(1 - P_{FA})}. \quad (3.60)$$

O débito máximo atingido pelo sistema é S_{sat}^* , calculado para J SUs e $P = P_{sat}^*$ obtendo-se a equação,

$$S_{sat}^* = (1 - P_U) \left(1 - \frac{1}{J}\right)^{J-1}. \quad (3.61)$$

Observa-se que quando o número de SUs é muito elevado, o desempenho se aproxima do valor assintótico do sistema *Slotted ALOHA* [Tan03], i.e. $\lim_{J \rightarrow \infty} = \frac{1-P_U}{e}$. No entanto, quando o número é baixo, é possível ter valores de desempenho bastante mais elevados, como está representado na figura 3.16. Os resultados da figura 3.16 foram obtidos através do simulador e das

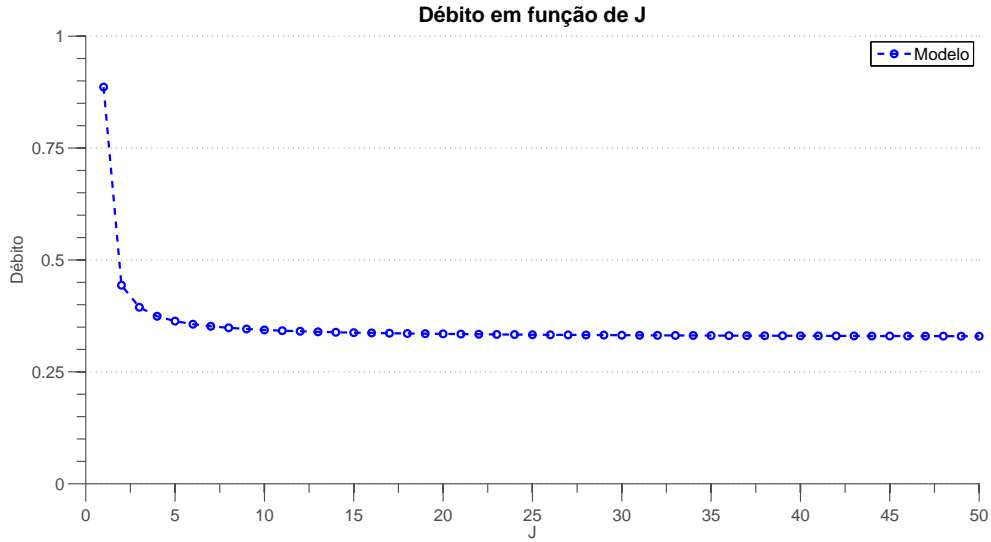


Figura 3.16: Resultados do débito máximo do sistema em função de J do modelo.

expressões do modelo utilizando a parametrização da tabela 3.5.

PU	N_S	N_T	γ	P	$\pi_{1,1}$	$\pi_{0,0}$	P_D	P_{FA}	λ
0.105644	42	425	77.812909	$\frac{1}{J}$	0.934448	0.992234	0.998345	0.001083	0.5

Tabela 3.5: Parâmetros de configuração da figura 3.16.

Por outro lado, quando o nível de carga do sistema se mantém abaixo do valor de S_{max} e é conhecido o número de SUs ($J\lambda < S_{sat}^*$), a rede atinge um valor de débito máximo que se mantém constante para um conjunto de valores de P , como é possível observar na figura 3.17, obtida usando os resultados do simulador e do modelo com os parâmetros da tabela 3.6. Este valor máximo é igual ao nível de carga que foi gerado pelos SUs do sistema. Ou seja, todo o tráfego que é gerado

pelos SUs é enviado.

PU	N_S	N_T	γ	J	$\pi_{1,1}$	$\pi_{0,0}$	P_D	P_{FA}	λ
0.502905	42	425	77.812909	3	0.936388	0.935747	0.997833	0.005993	0.04

Tabela 3.6: Parâmetros de configuração da figura 3.17 e 3.18.

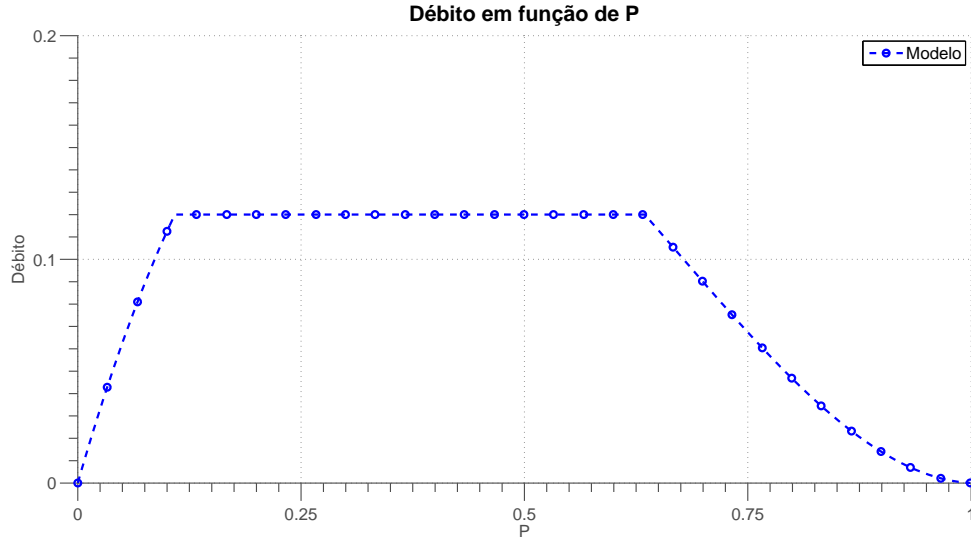


Figura 3.17: Débito médio medido com o modelo em função de P para $J\lambda < S^*$.

Adicionalmente, para $J\lambda < S_{sat}^*$, existe ainda um único valor de P , denominado por P^* , que minimiza o atraso médio de um pacote. O valor de P^* pode ser calculado através de:

$$\frac{\partial E[\Delta_T]}{\partial P} = 0 \quad (3.62)$$

Devido à complexidade da expressão 3.62, a sua derivada foi calculada por definição, através de um incremento de 10^{-7} e utilizou-se o método da bisseção para obter a sua solução.

A figura 3.18 representa a variação do atraso médio de um pacote em função de P . O mínimo do atraso médio é atingido quando $P = P^*$. A figura 3.18 foi obtida através dos resultados do simulador e do modelo com os parâmetros da tabela 3.6.

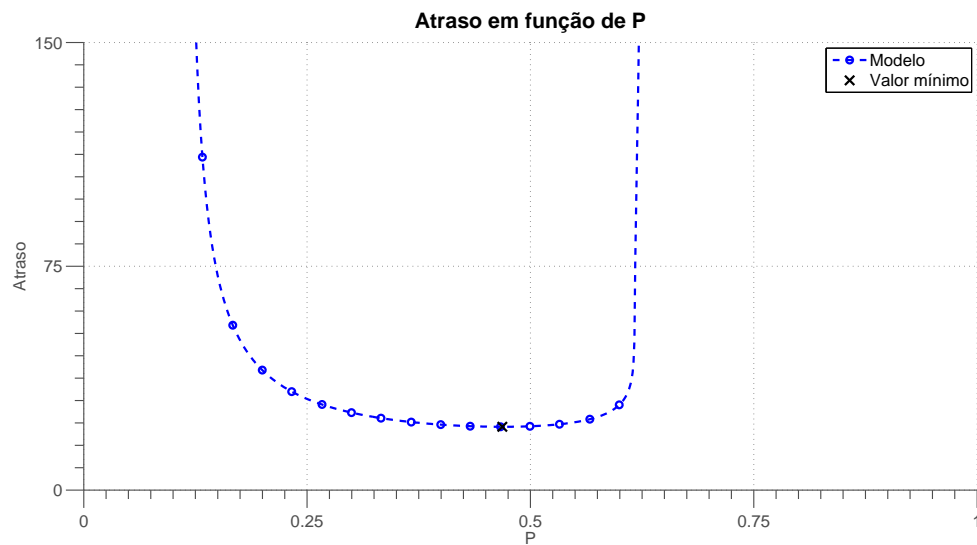


Figura 3.18: Atraso médio de um pacote medido pelo modelo em função de P para $J\lambda < S^*$.

Capítulo 4

Desempenho do sistema

Neste capítulo é apresentado o simulador do protocolo *p-persistent Slotted Aloha* adaptado para *RC* introduzido no capítulo anterior, elaborado utilizando o programa *Matlab*. Na secção 4.1 é apresentado o esquema geral do simulador, bem como a explicação detalhada dos seus módulos.

Este capítulo mostra também um conjunto de resultados do modelo e do simulador para o protocolo. Os resultados estão organizados em duas secções: na secção 4.2 é estudada a influência do parâmetro P no desempenho e o seu valor ótimo para um dado nível de carga; na secção 4.3 é estudado o desempenho do sistema perante diferentes níveis de carga.

4.1 Implementação do simulador

4.1.1 Modelo geral do simulador

Foi desenvolvido um simulador que modela o sistema introduzido no capítulo anterior através de uma sequência de eventos. A arquitetura do simulador desenvolvido está representada na Fig. 4.2. O simulador modela um *PU* e J *SUs* (transmissores) que utilizam o método de deteção de energia discutido no capítulo 2. Assume-se que quando não há colisão na transmissão, não existem erros na receção dos dados dos *SUs*. Os pacotes de dados são gerados no módulo de geração aleatória, no início da simulação, e são guardados numa estrutura de dados para serem utilizados ao longo do tempo de simulação. O tempo de simulação é gerido por um controlador de avanço de tempo. De notar que, quando é negada a transmissão de um pacote a um *SU*, o utilizador secundário mantém os seus pacotes de dados numa fila de espera de tamanho ilimitado para serem transmitidos assim que possível.

O simulador modela o efeito de captura de canal onde um utilizador de um meio partilhado “captura” o meio quando efetua a transmissão de dados, impossibilitando a comunicação de outros

utilizadores. Qualquer utilizador que pretenda aceder ao meio, após fazer o *sensing* e concluir que não existe nenhum utilizador a transmitir, acede ao canal com uma probabilidade P . No caso de dois ou mais utilizadores decidirem aceder ao meio ao mesmo tempo, ocorre uma colisão. Após a colisão, todos os utilizadores que tentaram transmitir falham a comunicação e, têm que efetuar uma retransmissão dos dados mais tarde.

O resultado final gerado pelo simulador são os valores do atraso médio por pacote, da probabilidade média das filas de espera estarem vazias e do débito.

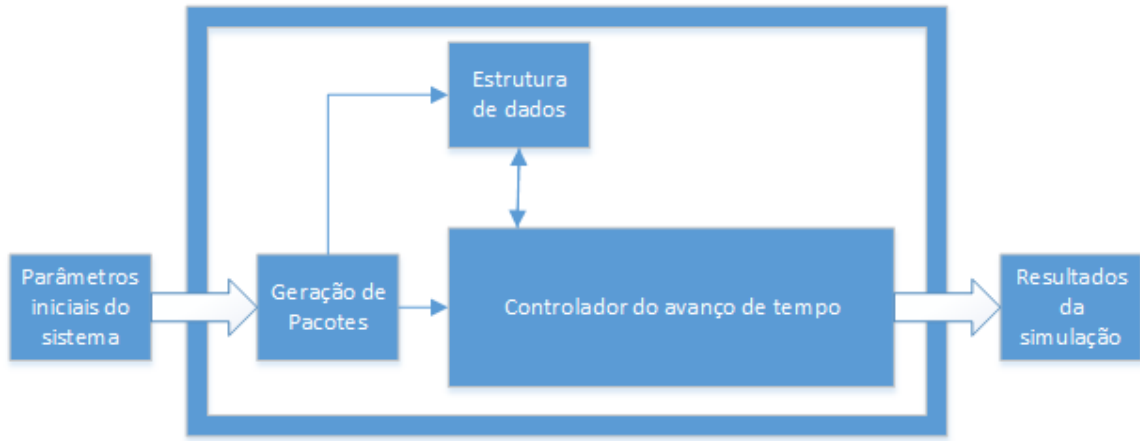


Figura 4.1: Módulos do simulador.

4.1.2 Parâmetros iniciais do sistema

Para configurar as especificações e parâmetros iniciais, o simulador carrega um ficheiro escolhido pelo utilizador com a especificação dos tempos de transição de atividade do *PU* que define a estatística dos tempos de transição. O simulador calcula a probabilidade do *PU* estar ativo, bem como as probabilidades do *PU* se manter ativo ou inativo no próximo *slot* de *SU*, $\pi_{1,1}$ ou $\pi_{0,0}$, respetivamente. Estas probabilidades são guardadas na estrutura de dados para mais tarde serem utilizadas juntamente com o modelo.

O simulador também oferece a possibilidade de configurar a frequência de amostragem e a percentagem do tempo de *sensing* dos *SUs*. Através da frequência de amostragem, é calculado o período de amostragem, também conhecido como o período de duração de um *slot* das tramas dos *SUs*. Para a simulação de todos os resultados desta tese, foi utilizado uma frequência de amostragem de 10 KHz e um período de duração de um *slot* de *SU* de 50 ms. Por outro lado, com a percentagem do tempo de *sensing*, são calculados os períodos de *sensing* e tempo de transmissão, mantendo-se o período total das tramas dos *SUs* constante nos 0.0213 segundos (N_T slots). De notar que se utilizou um valor de 10% da trama para o tempo de *sensing* pois, para a configuração

adotada nos testes, a partir deste valor, a interferência de um SU sobre um PU é menor do que 3% [LFO⁺13]. Além disso, com o período de tempo de trama e com o período de amostragem, é calculado o número de *slots* para cada fase do ciclo dos SUs (*sensing* e transmissão). O utilizador também configura o número de SUs (J) que irão estar ativos no sistema, assim como a taxa de chegada de pacotes pretendida por utilizador secundário, representada por λ .

Por fim, o simulador obtém o valor do limiar de energia necessário para os SUs detetarem a transmissão do PU , através do critério C_3 mencionado em [LFO⁺12]. Este critério tenta maximizar o sucesso de *sensing* por meio da maximização do produto entre P_D e a probabilidade de rejeição correta ($1 - P_{FA}$): $C_3 = P_D(1 - P_{FA})$. Por outras palavras, C_3 pretende aproximar P_D de 1 e P_{FA} de 0. Consequentemente, através das equações 4.1 e 4.2 de [LFO⁺12] e $1 \approx P_D(1 - P_{FA})$, o simulador obtém o valor do limiar de energia. De realçar que, de todos os critérios mencionados em [LFO⁺12], C_3 foi escolhido porque apresenta o melhor *tradeoff* entre interferência provocada em PU (garante quase proteção total do PU) e débito dos SUs .

$$P_D = Pr(M > \lambda_E | \mathcal{H}_1) = \mathcal{Q} \left(\frac{\gamma - (N_S + \lambda)}{\sqrt{2(N_S + 2\lambda)}} \right), \quad (4.1)$$

$$P_{FA} = Pr(M > \lambda_E | \mathcal{H}_0) = \mathcal{Q} \left(\frac{\gamma - N_S}{\sqrt{2N_S}} \right). \quad (4.2)$$

4.1.3 Geração aleatória da chegada de pacotes

De forma a validar o modelo analítico proposto no capítulo anterior, assume-se que no modelo a chegada de pacotes em cada nó é independente dos outros, sendo descrita através de uma distribuição de *Poisson*.

Para simular a geração de pacotes de dados às filas dos SUs , o simulador utiliza a expressão 4.3 para prever o intervalo de tempo entre a chegada de dois pacotes,

$$nextTime = -\frac{\ln(u)}{\lambda}, \quad (4.3)$$

sendo u um número aleatório com uma distribuição uniforme entre 0 e 1 e λ a taxa de chegada de pacotes.

Como a simulação começa num instante $T_0 = 0$ s, o simulador obtém o tempo de chegada do próximo pacote a partir deste instante T_0 . Subsequentemente, utiliza o instante T_1 , instante da chegada do primeiro pacote, para obter o tempo de aparecimento do próximo pacote. Este processo ocorre até ao tempo de chegada de um pacote ultrapassar o tempo total de simulação.

A partir dos tempos de intervalo da geração de pacotes, é possível obter o instante exato da

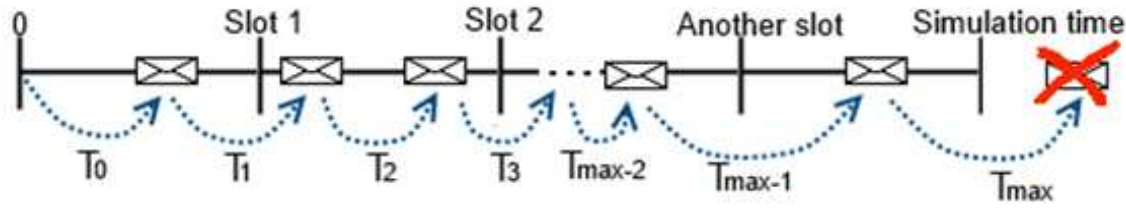


Figura 4.2: Geração de pacotes.

geração de cada pacote de dados, bem como o slot de chegada.

4.1.4 Controlador do avanço do tempo

O bloco principal do simulador é o controlador de avanço do tempo. Este bloco define quando o controlador avança de trama em trama de *SU*, desde o instante inicial T_0 do simulador até ao fim do tempo de simulação. Numa fase inicial de cada trama de *SU*, é modelado o sinal no meio durante os *slots* de *sensing*. Durante esta fase, o simulador analisa em que instante se dá uma alteração do sinal de *PU* (se for o caso) e, consoante o número de *slots* onde existe sinal do *PU* e ruído e de *slots* onde só existe ruído, é calculado o valor de energia presente no meio durante todo o período de *sensing* através das expressões das distribuições do ruído Gaussiano e sinal de *PU* definidas nas hipóteses representadas na equação (3.3). De notar que este valor é calculado através de uma distribuição normal, podendo os *SUs* obter valores diferentes entre si;

Em resultado dos sinais presentes do meio, podem ocorrer os seguintes eventos para cada utilizador secundário:

- **Evento de *sensing* do sinal de *PU*** - Cada *SU* compara o valor obtido na fase anterior com o *threshold* calculado no início da simulação e, se o valor for superior ao do *threshold*, o *SU* conclui que o *PU* está ativo. Caso contrário, o *SU* deteta apenas ruído no meio [LFO⁺13].
- **Processamento da atividade do *PU* em todos os *slots* da frame de *SU*** - O simulador determina a atividade do *PU* em todos os *slots* da trama em análise para concluir se o *PU* estava ativo/inativo e, de seguida, compara a sua conclusão com a análise de cada *SU*. No fim da simulação, as comparações anteriores são utilizadas para fins de cálculo das probabilidades de falso alarme e de deteção. Por exemplo, se o simulador observar que o *PU* está ativo numa determinada trama e se um *SU* conseguir realmente detetar o sinal do *PU* durante o seu período de *sensing*, afirma-se que o *SU* detetou corretamente o sinal de *PU*. Por outro lado, se o *SU* detetar uma transmissão do *PU* sem este estar ativo, é concluído que existiu um evento de falso alarme na trama;

- **Entrada dos pacotes recém-chegados na fila de espera de cada SU** - Após o processamento da atividade do *PU*, o simulador verifica na estrutura de dados se existe algum pacote gerado entre o final do período de *sensing* da trama anterior (instante da entrada de pacotes da trama anterior) e o final do período de *sensing* da trama atual. Todos os pacotes que chegarem durante este período de tempo são adicionados à fila de espera do respectivo *SU*;
- **Processamento dos tempos de atraso de cada pacote** - Este evento ocorre logo a seguir à entrada dos pacotes recém chegados na fila de espera, ou seja, no final do período de *sensing*. Durante esta fase, todos os pacotes dentro das filas de cada *SU* são analisados e recalculados os tempos de inatividade (*Vacation Time*), os tempos de espera e os tempos de serviço. Para cada pacote de dados, este processo de cálculo começa na trama em que estes são gerados, ocorre em todas as tramas e acaba no momento em que são transmitidos com sucesso.
- **Transmissão de pacotes** - Os *SUs* que não detetaram nenhuma transmissão do *PU*, acedem ao meio com uma probabilidade P . Se mais de um *SU* aceder ao meio, ocorre uma colisão e os pacotes têm de ser reenviados. Se apenas um *SU* aceder ao meio mas tiver ocorrido uma falha de deteção por parte do *SU*, também ocorreu uma colisão entre as transmissões do *PU* e do *SU*. É relevante indicar que, quando um utilizador captura o canal com o objetivo de enviar dados, se considera que a transmissão tem sempre a duração do período de transmissão completo.
- **Contador da fila de espera** - No final de cada ciclo do controlador, isto é, no final de cada trama dos *SUs*, são analisadas as filas de espera de todos os *SUs* e verificam-se quais as filas que estão vazias. No final da simulação, esta informação é usada para obter o valor da probabilidade das filas de espera estarem vazias P_{QE} e é obtida através de uma média por *slot* de *SU*. Existem dois modos de funcionamento do contador da fila de espera: análise da fila de espera apenas quando *PU* está inativo (utilizado para os resultados da primeira e segunda abordagens do capítulo 3) e análise da fila de espera a tempo inteiro (*PU* ativo e inativo) (utilizado na abordagem final do mesmo capítulo).

4.1.5 Resultados da simulação

Durante a simulação, para todas as tramas dos *SUs*, vai sendo guardada numa estrutura de dados a informação necessária para que se possam gerar os resultados. São armazenados o tamanho das filas de espera dos *SUs* por trama, se os utilizadores secundários detetaram ou não

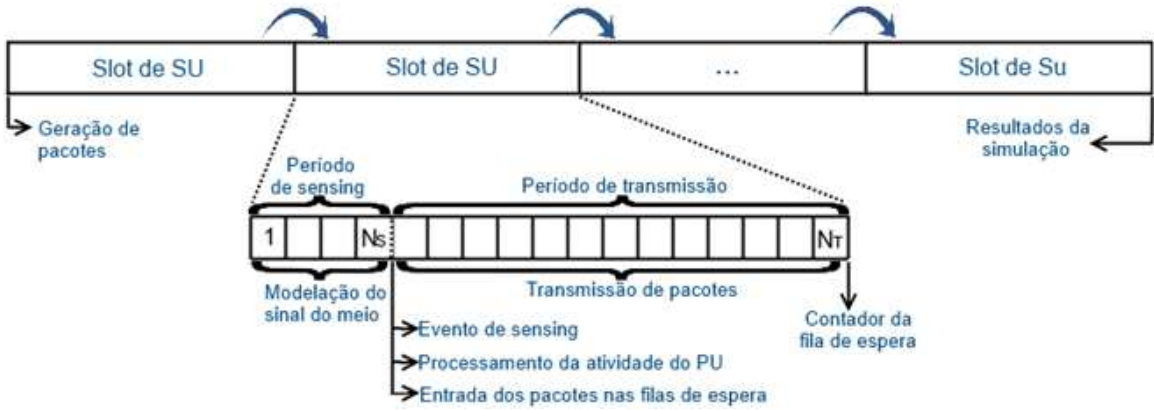


Figura 4.3: Exemplo do controlador de tempo e ilustração dos instantes em que ocorrem os eventos.

a atividade do PU , se transmitiram ou se as filas estavam vazias, o instante da chegada, quando são enviados e os tempos de inatividade, espera e serviço por pacote. Tal como já foi dito no início do capítulo, no fim da simulação são calculados os resultados finais da simulação. Por sua vez, quando se efetua uma simulação, obtém-se os resultados acima mencionados para um nível de carga λ . Portanto, para se obter informação suficiente para construir um gráfico, é necessário correr o simulador diversas vezes com os mesmos parâmetros mas variando os valores de λ . Por outro lado, para adquirir os resultados da otimização do sistema e do P ótimo, é necessário manter um valor de λ constante mas variar o valor de P .

4.2 Análise da variação de P

Nesta secção, é estudada a influência do parâmetro P no desempenho do sistema, identificando-se o seu valor ótimo para um dado nível de carga. São comparados os valores obtidos com o modelo analítico e com o simulador.

Os resultados de todas as figuras desta secção foram obtidos com as seguintes parametrizações:

P_U	N_s	N_T	$\pi_{1,1}$	$\pi_{0,0}$	N_S	γ	P_D	P_{FA}
0.502905	42	425	0.936388	0.935747	42	77.815435	0.997827	0.005997

Tabela 4.1: Parametrização de configuração da variação de P .

4.2.1 Análise de P_{QE}

A figura 4.4 mostra os valores de P_{QE} do modelo e do simulador em função de P para $J = 2$ com vários valores de carga, enquanto a figura 4.5 mostra o comportamento para $J = 3$, também com dois valores de carga. Tal como foi observado no capítulo anterior, o modelo subestima o valor de P_{QE} medido com o simulador. Como esperado, verifica-se que a subida da taxa de chegada de pacotes reduz o valor de P_{QE} para o mesmo valor de P , reduzindo o intervalo de valores de P onde o sistema não está saturado. No entanto, a curva teórica acompanha o andamento do valor medido com o simulador, permitindo estudar os valores extremos.

Tanto na figura 4.4 como na figura 4.5 é possível verificar que existe um valor de P onde o sistema alcança o melhor desempenho possível para uma configuração específica, com o valor máximo de P_{QE} . Este valor é denominado P ótimo, e representado com P^* . Nas duas figuras, é visível que o valor de P^* na simulação é quase idêntico ao do modelo, validando-o.

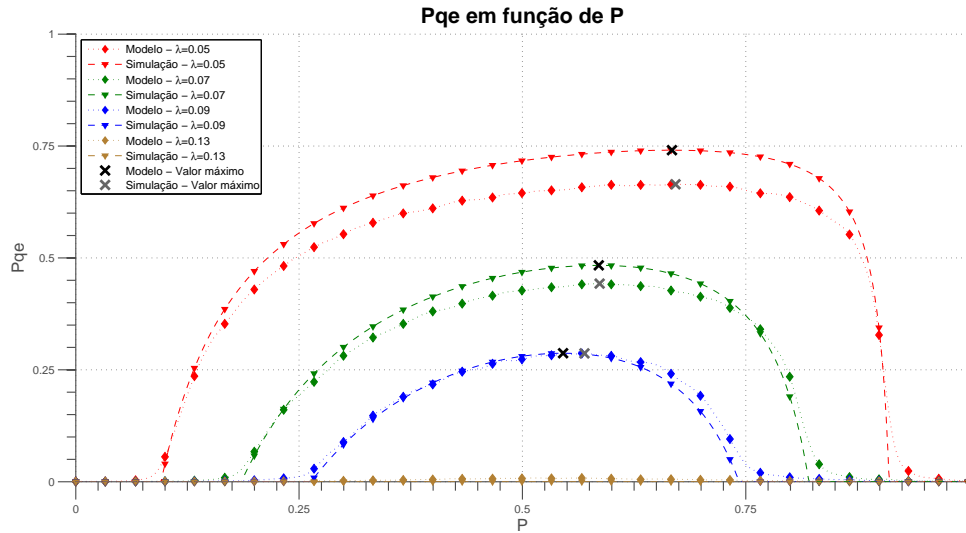


Figura 4.4: P_{QE} em função de P . Impacto da variação de λ nos valores de P_{QE} para $J = 2$.

Na figura 4.5, também se observam dois cenários diferentes. Nos primeiros dois resultados é possível verificar que o sistema não se encontra em saturação (azul e vermelho). Por outro lado, nos restantes (verde e dourado), o sistema encontra-se num cenário de saturação, onde o P_{QE} é sempre nulo.

4.2.2 Análise do débito

Analisando agora o débito do sistema através das figuras 4.6 e 4.7, com as mesmas configurações das duas figuras anteriores, verifica-se que os resultados analíticos do modelo são valida-

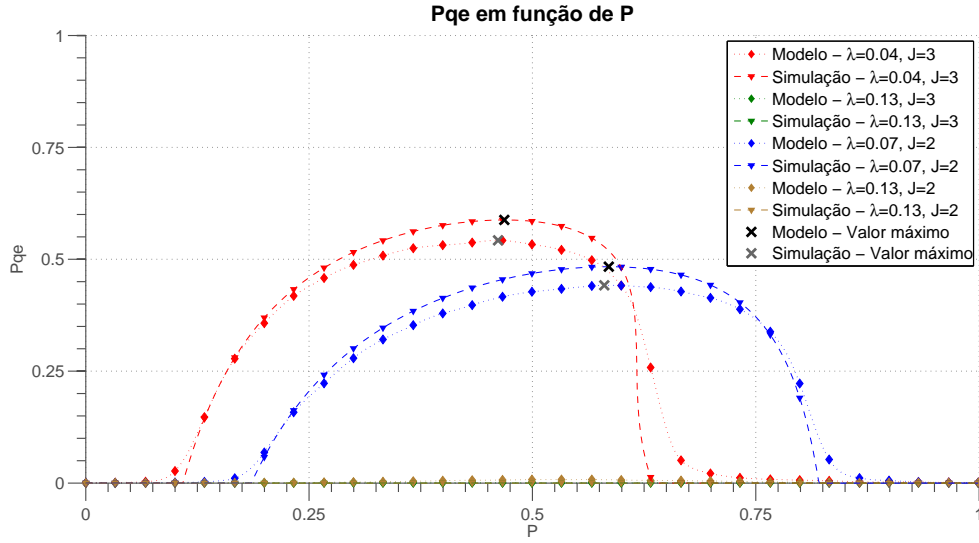


Figura 4.5: P_{QE} em função de P . Impacto da saturação do sistema nos valores de P_{QE} . Resultados vermelho e verde referentes a $J = 3$. Os restantes foram simulados para $J = 2$.

dos pelas simulações, onde se constata que os valores do modelo seguem precisamente os resultados da simulação.

O aumento do λ tem uma influência negativa no sistema. No caso do débito, o crescimento de λ causa um aumento do débito total máximo do sistema, no entanto, reduz o intervalo de valores de P^* para se alcançar o pico do mesmo. Por sua vez, se o λ for suficientemente grande para saturar o sistema, é possível atingir o débito máximo total do sistema para a respetiva parametrização, juntamente com um único P^* .

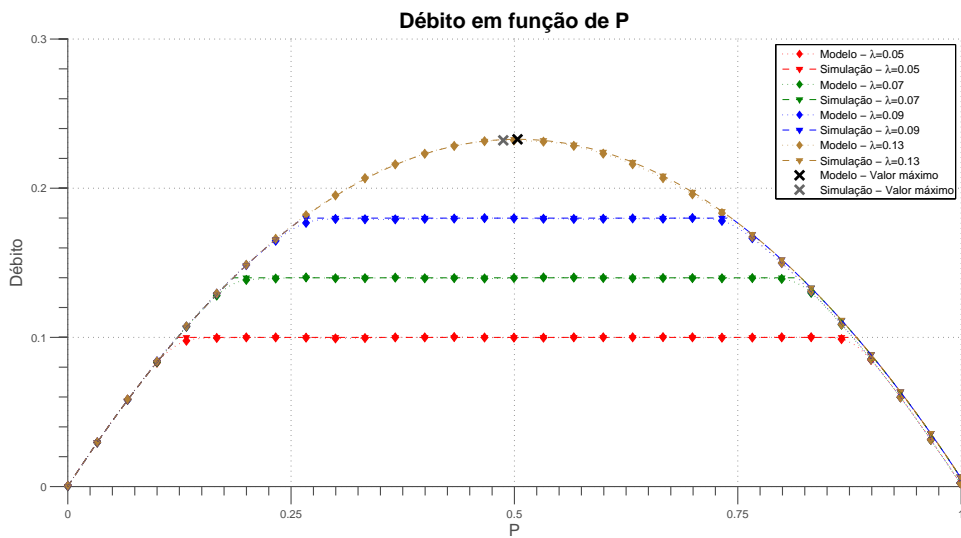


Figura 4.6: Débito em função de P . Impacto da variação de λ nos valores do débito para $J = 2$.

É relevante salientar que os P^* s obtidos através do modelo e do simulador não são os mesmos porque seria necessário despende de elevados períodos de tempo na elaboração da simulação para que o P^* da simulação se tornasse muito preciso. No entanto, o tempo consumido na simulação é suficiente para validar os resultados do modelo com rigor. Apesar de tudo, os P^* s obtidos encontram-se próximos um do outro.

Por outro lado, na figura 4.7 são demonstradas duas parametrizações diferentes do sistema (apenas diferem em J) onde, em cada uma delas, são exibidos os cenários de saturação e não saturação da rede. Em ambos os casos, nos resultados de saturação, observa-se que o débito sobe à medida que se aumenta o P até atingir o débito máximo total no ponto P_{sat}^* . Com o contínuo aumento de P a partir de P_{sat}^* , o débito total sofre uma descida gradual até atingir o valor 0 em $P = 1$. Por outro lado, no caso da não saturação do sistema, verifica-se que existe um intervalo de valores de P onde o débito é igual ao valor de pacotes gerados no sistema.

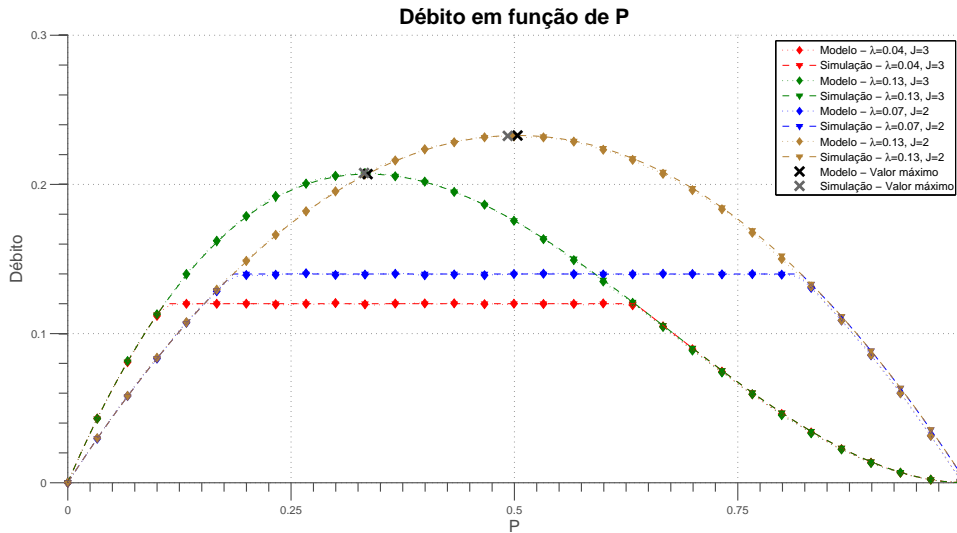


Figura 4.7: Débito em função de P . Impacto da saturação do sistema nos valores do débito.

4.2.3 Análise do atraso médio de pacotes

As figuras 4.8 e 4.9 mostram os valores obtidos com o modelo e com o simulador para o atraso médio de pacotes para os dois cenários analisados nesta secção. Apesar de existir uma diferença entre os valores teóricos e simulados, a curva teórica acompanha o andamento do valor medido com o simulador, permitindo estudar os valores extremos, tal como em P_{QE} .

Comparando a figura 4.8 com os resultados de P_{QE} (figura 4.4), é possível observar que, para uma carga baixa, o atraso médio decresce enquanto P_{QE} aumenta. Conforme o P_{QE} diminui, constata-se uma subida progressiva do atraso médio.

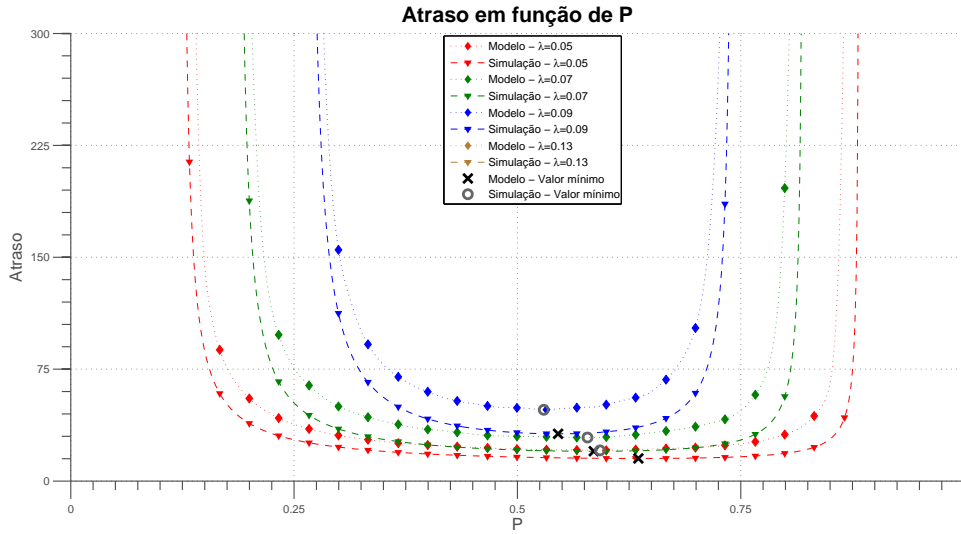


Figura 4.8: Atrazo médio em função de P . Impacto da variação de λ nos valores do atraso médio para $J = 2$.

É importante realçar que os valores de P onde o sistema alcança o resultado de P_{QE} máximo correspondem aos valores de P onde o atraso médio é mais pequeno, como pode ser constatado nas figuras 4.9 e 4.5.

Por fim, na figura 4.8 ilustra-se o impacto de λ no sistema. Observa-se que, quanto maior é o valor da taxa de chegada de pacotes, maior é o valor do atraso médio de pacotes para todos os valores de P . De notar que, para os respetivos parâmetros do sistema e $\lambda = 0.13$ pacotes/slot, o sistema encontra-se saturado, pelo que atinge valores de atraso médio de um pacote tão elevado que nem está presente na figura.

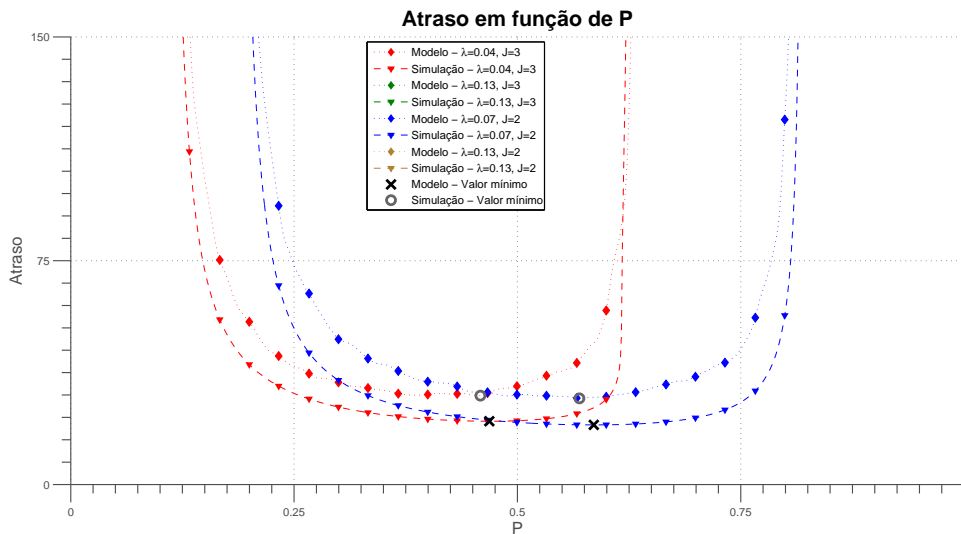


Figura 4.9: Atrazo em função de P . Impacto da saturação do sistema nos valores do atraso médio.

4.3 Análise do desempenho

Nesta secção, é analisado o desempenho do sistema em função do número de SUs e da probabilidade do PU estar ativo.

4.3.1 Escalabilidade do sistema

As figuras desta secção demonstram o impacto do aumento do número de SUs ativos e a escalabilidade do desempenho deste sistema. Os resultados das figuras desta secção foram obtidos com as parametrizações da tabela 4.2. O parâmetro P foi definido como $P = \frac{1}{J} \simeq P_{sat}^*$. Desta forma, só é necessário ter conhecimento do valor de P para a simulação dos resultados.

P_U	N_S	N_T	$\pi_{1,1}$	$\pi_{0,0}$	P	N_S	γ	P_D	P_{FA}
0.502565	42	425	0.935762	0.935209	$\frac{1}{J}$	42	77.817634	0.997642	0.006173

Tabela 4.2: Parametrização de configuração da escalabilidade do sistema.

A figura 4.10 representa P_{QE} em função da carga. Pode-se observar que a probabilidade da fila de espera estar vazia P_{QE} tende a diminuir mais rapidamente com um aumento de carga para um grande número de SUs . Além disso, a figura mostra que este modelo analítico é um pouco pessimista face à determinação do limite de saturação, visto que este ocorre só se verifica para grandes cargas.

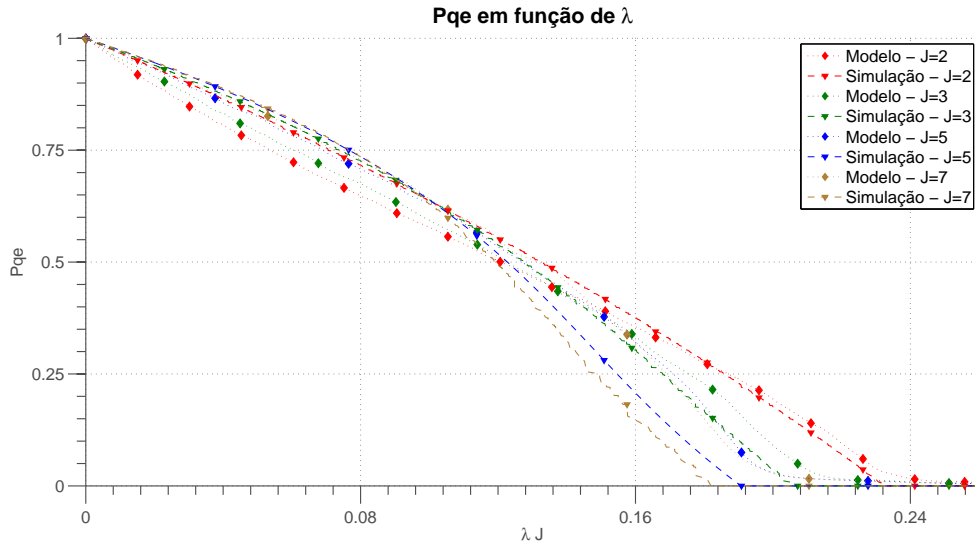


Figura 4.10: P_{QE} em função de λJ . Impacto da variação de J s nos valores de P_{QE} .

Por outro lado, na figura 4.11 é possível verificar que enquanto se aumenta a taxa de chegada de pacotes, o número de dados enviados com sucesso tende a crescer até atingir um valor máximo

chamado de débito total de saturação. O débito máximo é atingido quando o sistema satura. Isto acontece porque o número de pacotes que entram na fila de espera é maior do que o número de pacotes que são transmitidos com sucesso, ou seja, as filas de espera nunca estão vazias e, por isso, os *SUs* decidem transmitir sempre que o *PU* não está ativo.

O aumento do número de *SUs* ativos na rede influencia negativamente o débito total do sistema. Quanto maior for o número de *SUs* ativos na rede, menor será o débito total do sistema até que este atinja o seu valor limite ($\lim_{J \rightarrow \infty} = \frac{1-P_U}{e} = 0.182996$ pacotes/slot, com $P_U = 0.502565$).

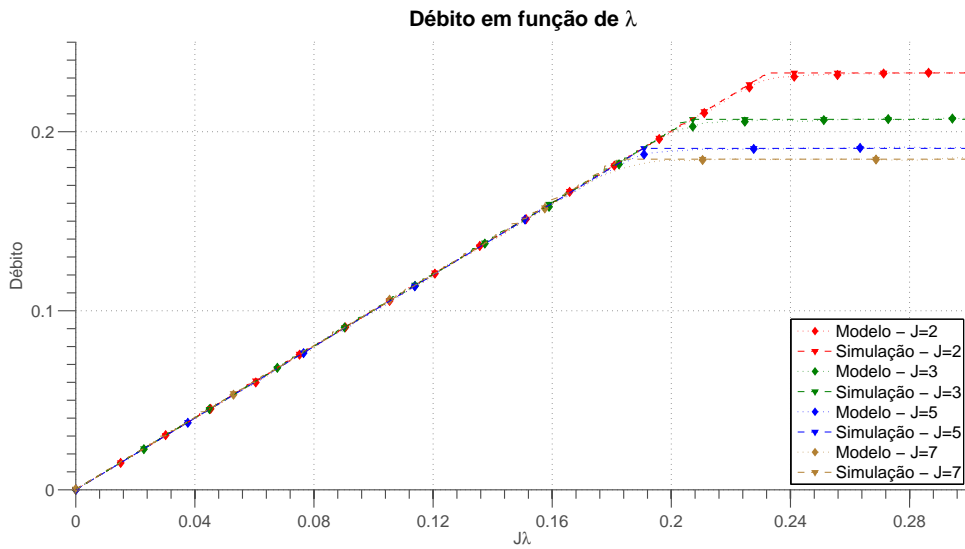


Figura 4.11: Débito em função de $J\lambda$. Impacto da variação de J nos valores do débito.

Na figura 4.12 verifica-se que, quando a taxa de chegada de pacotes cresce, o atraso médio de transmissão dos pacotes acompanha este crescimento. No entanto, quando o sistema está perto do limite de saturação, os tempos do atraso começam a agravar-se e observa-se uma acentuação abrupta na demora da transmissão dos pacotes. Adicionalmente, a diferença entre os valores do modelo e do simulador acentua-se nesta zona devido à imprevisibilidade do comportamento do sistema saturado.

Mais uma vez, verifica-se que o aumento do número de *SUs* ativos no sistema prejudica significativamente o desempenho do sistema. Quanto maior é o número de *SUs*, menor é a carga necessária de pacotes para se atingir a zona de aumento rápido do atraso médio.

Por fim, confirma-se, novamente, a validação do débito e do atraso médio de pacotes pois os resultados do simulador seguem os valores obtidos através da simulação.

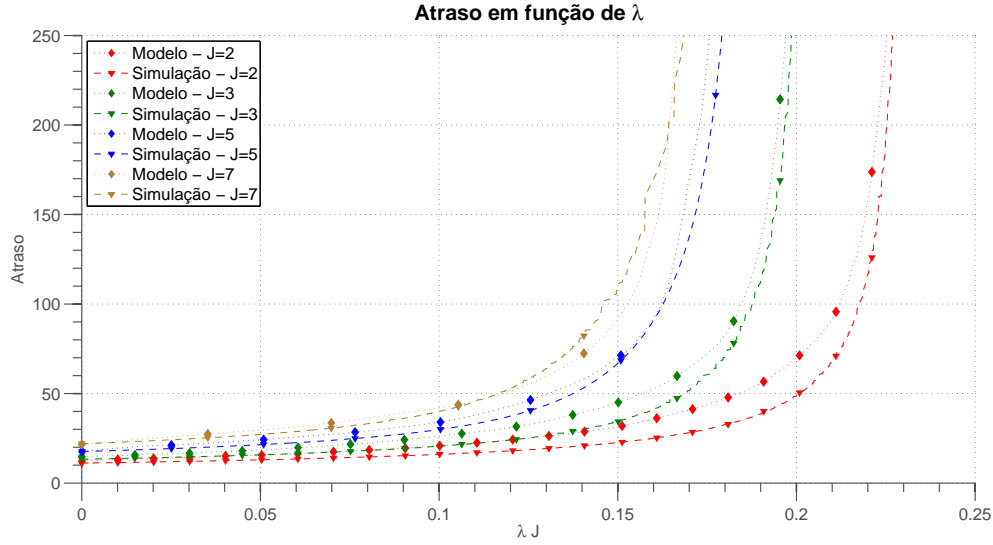


Figura 4.12: Atraso médio em função de λJ . Impacto da variação de J s nos valores do atraso médio.

4.3.2 Impacto do P_U no sistema

Os resultados das figuras que se encontram nesta secção foram obtidos através da modelação e simulação com os seguintes parâmetros:

N_S	N_T	J	P	γ	P_D	P_{FA}
42	425	3	$\frac{1}{3}$	77.812909	0.998342	0.001084

Tabela 4.3: Parametrização de configuração do impacto do P_U no sistema.

Analisando o desempenho do sistema em função da probabilidade do PU estar ativo, observa-se que o aumento do tempo de atividade do PU tem consequências negativas no desempenho dos SUs . Tal facto deve-se à necessidade que os SUs têm de utilizar apenas os períodos de inatividade do PU . Com o aumento do período de atividade do PU , as oportunidades de transmissão dos SUs tornam-se mais escassas. Desta forma, verifica-se na figura 4.13 que, com o aumento deste período, o sistema atinge a zona de saturação com uma taxa de chegada de pacotes mais baixa. Na figura 4.14, é possível observar que com a redução do período de atividade do PU o débito total máximo do sistema aumenta, pois o número de possíveis tramas de transmissão para os SUs também é maior. Mais uma vez, o sistema alcança o débito máximo total e o ponto de saturação no mesmo valor de λ em que P_{QE} atinge o valor 0.

Na figura 4.15, tal como na figura 4.12, o atraso médio toma um crescimento acelerado antes do ponto de saturação e o aumento da probabilidade do PU estar ativo tem consequências negativas para o sistema, provocando uma amplificação do atraso médio dos pacotes.

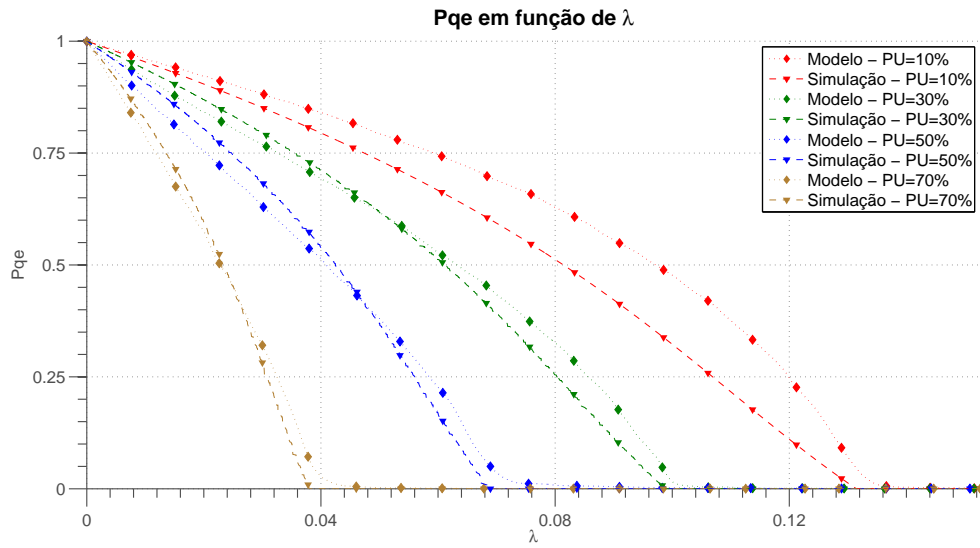


Figura 4.13: P_{QE} em função de λ . Impacto da variação de P_U nos valores de P_{QE} .

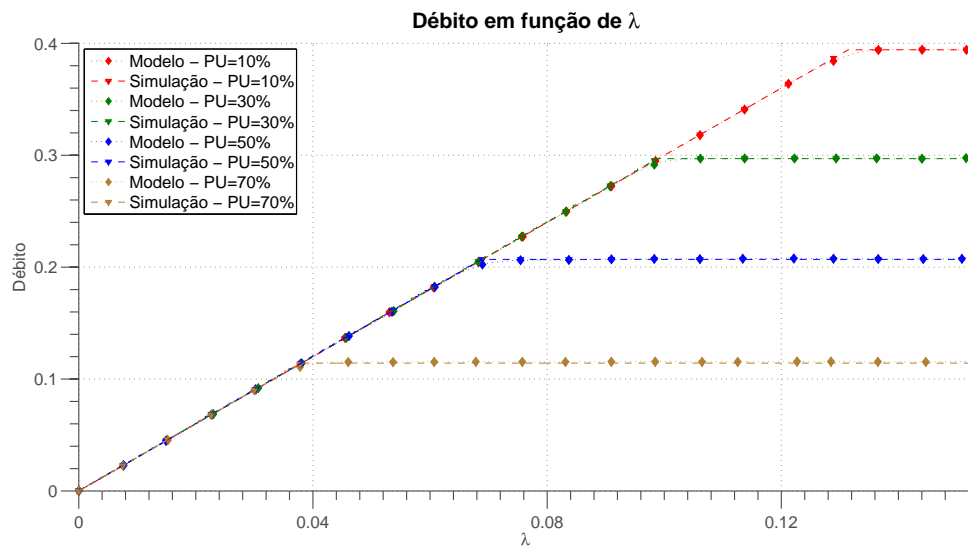


Figura 4.14: Débito em função de λ . Impacto da variação de P_U nos valores do débito.

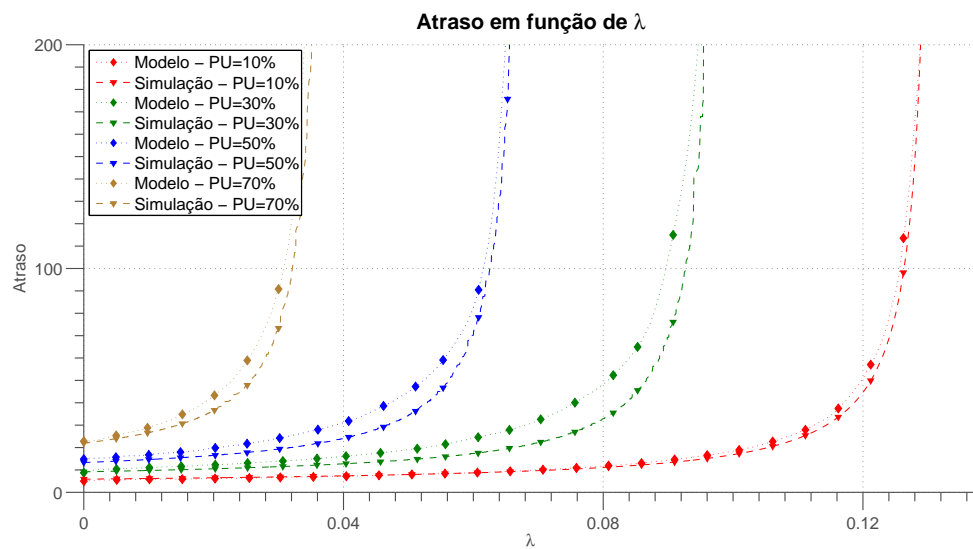


Figura 4.15: Atraso médio em função de λ . Impacto da variação de P_U nos valores do atraso médio.

Capítulo 5

Conclusões

5.1 Considerações Finais

Neste capítulo são apresentadas as considerações finais sobre o documento realizado, assim como futuros desenvolvimentos no modelo proposto.

Neste documento foram abordados os temas das redes de *RC*. Este trabalho foca principalmente a análise dos mecanismos de controlo de acesso ao meio utilizados nas redes de *RC*.

Numa primeira fase foi modelado o comportamento dos utilizadores secundários numa rede de rádio cognitivo utilizando o protocolo *Slotted ALOHA P*-persistente. Foi realizado um simulador do sistema e proposto um modelo estocástico num processo de refinamento sucessivo, onde se procurou ajustar o valor do débito do modelo com o valor medido no simulador. Foi modelada a probabilidade média em substituição do valor real desta probabilidade, variável ao longo do tempo. Por outro lado, o valor de P_{QE} foi usado para estimar o atraso médio de pacotes, tendo-se observado um ligeiro desvio, que se intensificou antes do modelo entrar na zona de saturação devido ao comportamento mais instável do sistema quando a fila de espera tem grandes variações de pacotes. Quando P_{QE} está com um valor muito alto, valor 1 ou com um valor nulo, a variação da probabilidade da fila de espera estar vazia é baixa, pelo que os valores do modelo têm um menor desvio face aos resultados da simulação, validando com sucesso os valores do atraso médio por pacote. Também é observado que os resultados do débito são perfeitamente validados. Assim, apesar das pequenas variações em P_{QE} e no atraso médio, conclui-se que o modelo foi validado através dos resultados das simulações com sucesso.

Numa segunda abordagem, procedeu-se à análise do desempenho do sistema para diferentes configurações. Durante esta análise, verificou-se que existe um valor ótimo da probabilidade de um *SU* aceder ao meio (P^*). Provou-se analiticamente que, através da utilização deste valor, o

sistema atinge o valor máximo do débito. É de salientar que, para maximizar o débito, o sistema exige uma taxa de chegada de pacotes suficientemente alta para atingir a saturação. Por outro lado, verificou-se experimentalmente que o valor de P_{QE} do sistema atinge o seu máximo no mesmo valor de P em quem o atraso médio de pacotes do sistema é mínimo.

Também é analisado o desempenho da rede quanto à escalabilidade da rede secundária e quanto à atividade da rede primária. Assim, conclui-se que o débito total da rede secundária decresce à medida que se aumenta o número de utilizadores nesta rede. Por outro lado, quando a atividade do utilizador primário cresce, existe um decréscimo de oportunidades de acesso no espetro para serem utilizados pela rede secundária. Isto leva a que o débito total da rede secundária decresça também. O aumento do tempo de atividade da rede primária também influencia negativamente o atraso médio por pacote da rede secundária: quanto mais ativo for um PU , mais rapidamente se atinge valores de atraso incrivelmente altos.

Como conclusão final, o estudo e modelação da otimização da contenção no acesso ao meio num protocolo deste tipo numa rede de RC pode originar uma melhor eficácia no desempenho, nomeadamente em termos de débito e atraso.

5.2 Trabalho Futuro

O protocolo proposto no capítulo 3 é baseado num modelo de colisão. Sob este modelo, a capacidade de uma rede sem fios é limitada principalmente pelas transmissões simultâneas de pacotes. As transmissões simultâneas levam ao aparecimento de colisões e a uma degradação significativa do débito total da rede. Adicionalmente, as retransmissões pioram muitas vezes a situação. Todas estas desvantagens estão relacionadas com a receção de apenas um pacote de cada vez nos recetores (*Single Packet Reception (SPR)*). Posto isto, para trabalho futuro, propõe-se a utilização de técnicas de *Multi Packet Reception (MPR)*, solucionando o problema do limite de débito provocado pela utilização de um modelo de acesso único ao canal (modelo de colisões) através de métodos de acesso aleatório ao meio como o *Network-assisted Diversity Multiple Access (NDMA)*. No *NDMA*, as retransmissões controladas pelo protocolo são utilizadas para resolver as colisões via separação no recetor.

Anexos

Apêndice A

Publicações

Performance of a Cognitive p -persistent Slotted Aloha Protocol

J. Reis(*), M. Luís(*, †), L. Bernardo(*, †), R. Oliveira(*, †), R. Dinis(*, †), P. Pinto(*, †)

(*) CTS, Uninova, Dep.^o de Eng.^a Electrotécnica, Faculdade de Ciências e Tecnologia, FCT,

Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

(†) Instituto de Telecomunicações, Lisboa, Portugal

Abstract—This paper proposes a new analytical model for the performance of a p persistent slotted Aloha medium access protocol in a Cognitive Radio Network (CRN) composed by a single Primary user (PU) and several Secondary users (SUs) with Poisson traffic. SUs run a synchronized sensing-access operation cycle, and when they have a packet to transmit and the channel is sensed free, they access to the data slots with probability p . PU change their activity state (ON/OFF) independently of the SUs, and an energy-based sensing performance model is considered. This paper proposes a new analytical model for the SU's delay and throughput which, contrarily to the existing models, considers the effect of the duration of the PU's ON state in the SUs delay. The model is validated through simulations, and the optimal p value is calculated for different scenarios. Two context-aware configurations are proposed: the optimal p value when only the number of SUs is known, and when the number and load are known.¹

Index Terms—Cognitive Radio; p -persistent Slotted Aloha; Analytical Performance Evaluation; Context-aware Configuration.

I. INTRODUCTION

The performance of a non-licensed secondary user (SU) running random access medium access control (MAC) protocols in a single radio Cognitive Radio Network (CRN) is limited by the impossibility of sensing and transmitting simultaneously. Due to this limitation, SUs usually adopt an operation cycle where sensing and transmission operations occur consecutively. For a single channel CRN, SUs start to sense the spectrum during a synchronized fixed amount of time (sensing period), where no SU transmits. Depending on the output of the sensing, if it does not detect the activity of licensed primary users (PUs), the SUs can transmit during a fixed amount of time (transmission period). However, the transmission is only successful when only one SU transmits. The transmission period may be lost due to collisions (when more than one SU transmits), or may be idle.

The slotted Cognitive Radio-Carrier Sense Multiple Access (CR-CSMA) [1] and slotted CR-Aloha [2] adapted respectively the standard CSMA and Slotted Aloha protocols to guarantee the PU's protection. As described above, all SUs synchronize the sensing operation according to a frame structure, and access the data slots using the CSMA or Aloha

protocols. They back off for a random number of data slots (between 0 and a maximum window size) when a transmission fail. CR-CSMA and CR-Aloha differ when the SU finds the channel occupied by the PU; CR-CSMA waits until the medium becomes empty (including the data slots), while CR-Aloha waits a random time only when it is occupied by the PU. CR-CSMA/CA (Collision Avoidance) [3] proposes an alternative approach: it lets nodes run asynchronously, but locks the SU's channel access before the sensing operation using a Prepare-To-Sense (PTS) control frame. Thereon it runs the Ready-To-Send/Clear-To-Send frame exchange of the conventional CSMA/CA protocol. In all cases, these papers analysed mainly the influence of the sensing parameters in the system performance, but either ignored the influence of the back off window used (for CR-CSMA and CR-Aloha), or the effect of the hidden-node problem (CR-CSMA/CA). p -persistent Slotted aloha was analysed in [4], but the duration of the PU active state was not considered in the analysis. When the p value approaches one, it is not valid the assumption that the probability of finding the channel occupied by the PU is assumed constant and independent in two consecutive slots.

In this paper we propose an alternative analytical model for the performance of a unsaturated and saturated p -persistent Slotted CR Aloha protocol. It models the throughput and delay, considering the performance of the energy sensing based approach presented in [5]. The model also addresses the problem of using context aware information, like the number of SUs and the sensing performance estimation, to define the optimal access probability parameter p .

The system overview, the PU behaviour and the sensing model is presented in section II. The system's performance is analyzed in section III and its optimization is approached in section IV. A set of performance results is presented in section V and section VI contains the conclusions.

II. SYSTEM CHARACTERIZATION

We consider that a single PU may access the channel and multiple SUs may access in an opportunistic way, such is defined in [5]. SUs' operation cycle includes the sensing and transmission periods, which facilitates the synchronization of the sensing task. Sensing and transmission period durations are represented by T_S^{SU} and T_D^{SU} respectively, as illustrated in Figure 1. The SU's frame, $T^{SU} = T_S^{SU} + T_D^{SU}$, contains N_T slots, where each slot duration is given by the channel

¹This work was supported by the FCT/MEC projects MANY2COMWIN EXPL/EEI-TEL/0969/2013; ADIN PTDC/EEI-TEL/2990/2012; IT UID/EEA/50008/2013; Femtocells PTDC/EEA-TEL/120666/2010; and COPWIN PTDC/EEI-TEL/1417/2012.

sampling period adopted in the spectrum sensing task. The first N_S slots define the sensing period duration, and the remaining ones ($N_S + 1$ to N_T) represent the transmission period duration. It is assumed that all SUs are synchronized.

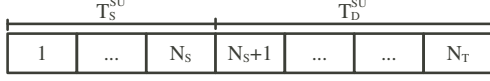


Fig. 1. SU's frame structure representing SU's operation cycle.

The PU's active period duration and inactive period duration are denoted by T_F^{PU} and T_F^{PU} , respectively. The SUs cycle length ratio, δ , is defined by the ratio between the PU shortest period duration and the SUs cycle period:

$$\delta = \frac{\min(T_F^{PU}, T_F^{PU})}{T^{SU}}. \quad (1)$$

Regarding the interference model, it is assumed a worst case scenario where a PU is unable to decode a received frame when it overlaps in time with a SU transmission.

A. Primary user model

To model the PU's activity, we consider the periodic channel sampling process described above, which is modelled by the Markov chain illustrated in Figure 2. Considering the stationary behavior of a PU during a SU's frame and $T_F^{PU} < T_F^{PU}$, the probabilities of a PU changing its behavior to active and inactive, π_{01} and π_{10} , respectively, or maintaining its behavior inactive or active, π_{00} or π_{11} , are given by [5]:

$$\begin{aligned} \pi_{10} &= \frac{1}{\delta N_T}, & \pi_{11} &= 1 - \pi_{10}, \\ \pi_{01} &= \frac{P_\beta}{\delta N_T(1 - P_\beta)}, & \pi_{00} &= 1 - \pi_{01}, \end{aligned} \quad (2)$$

where P_β represents the probability of a PU node staying ON. We assume that the PU does not change twice its state during a SU slot time.

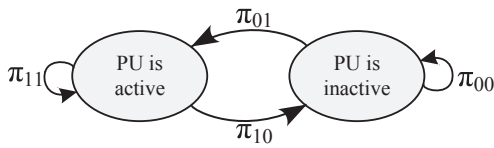


Fig. 2. A two state birth-death process describing the PU activity model.

B. Sensing model

PU nodes may change their state at any intermediate slot of a SU frame. In this paper we consider that all packets that are transmitted in a slot where a PU state transition occurs are not received correctly; this is a pessimistic approach, given that packets can be successfully received with the PER calculated in [5].

During the detection stage, each SU calculates the amount of energy received in the N_S samples $Y = \sum_{k=1}^{N_S} |x(k)|^2$, being then compared with the energy threshold γ to decide whether a PU is present or absent. $s(k)$ denotes the signals

transmitted by the PUs, with distribution $\mathcal{N}(\mu_s, \sigma_s^2)$. Y can be approximated to a Gaussian distribution [5] when the PU is OFF ($\mathcal{N}(N_S, 2N_S)$ for \mathcal{H}_{00}) or ON ($\mathcal{N}(N_S + \lambda, 2(N_S + 2\lambda))$ for \mathcal{H}_{11}), where $\lambda = \sum_{k=1}^{N_S} (\mu_s / (1 + \sigma_s))^2$ represents the sum of N_s samples of SNR collected from the channel during the sensing period. For a single PU, the probability of detection P_D and probability of false alarm P_{FA} are represented by

$$P_D = Pr(Y > \gamma | \mathcal{H}_{11}) = \mathcal{Q} \left(\frac{\gamma - (N_S + \lambda)}{\sqrt{2(N_S + 2\lambda)}} \right), \quad (3)$$

$$P_{FA} = Pr(Y > \gamma | \mathcal{H}_{00}) = \mathcal{Q} \left(\frac{\gamma - N_S}{\sqrt{2N_S}} \right), \quad (4)$$

where $\mathcal{Q}(\cdot)$ is the complementary distribution function of the standard Gaussian.

III. ANALYTICAL PERFORMANCE MODEL

This section analyses the performance of the CR p -persistent Aloha for a system with J SUs with infinite packet retransmissions, considering Poisson traffic load.

The SU's behaviour can be modelled by a sequence of relevant epochs, in which the PU is idle and the SUs may transmit with success, and irrelevant epochs, in which the PU is active or is changing its state. We are not considering any successful SU transmission that might occur during a SU frame where the PU changes its state from active to idle during the sensing slots of the SU frame [5].

A SU transmits in a slot of a relevant epoch with probability p when it has a packet in its queue, and it remains inactive when it does not have any packets in the queue. The following analysis considers a probability of a SU having its queue empty, denoted by ϱ , and uses it to estimate the number of SUs transmitting in each slot of the relevant epochs. In practice, the number of packets in the SU's queue is a stochastic process that changes in time - its expected value decreases in time after a sequence of relevant epochs and increases during the irrelevant epochs. Therefore, we are approximating the time variant stochastic process by a stationary process with the same average.

A. SU Success Probability

The fading and other channel effects may be hidden when the signal-to-noise ratio is set high enough. In these conditions, when the PU is idle, the SU transmission failures result only from collisions with other SUs. Thus a packet transmission is successful when only one single SU transmits during a relevant slot. Therefore, the error probability for a packet transmitted for the first time when the PU is sensed idle, ϵ_1 , is given by

$$\epsilon_1 = 1 - (1 - p(1 - P_{FA})(1 - \varrho))^{J-1}. \quad (5)$$

The success probability for the first transmission of a packet given that there is a packet in the queue, ς_1 , takes into account the SU's transmission probability,

$$\varsigma_1 = p(1 - P_{FA})(1 - \epsilon_1). \quad (6)$$

When a packet fails its initial transmission, there is at least one additional SU with packets in the queue. Assuming that SU queues are independent and identically distributed (i.i.d.), the new error probability will depend on the distribution of the number of SUs with packets to transmit, denoted by ζ . The probability mass function of the number of SUs with non-empty queues, ϕ_k , is

$$\begin{aligned}\phi_k &= \mathbb{P}\{\zeta = k \mid \epsilon_1\} = \frac{\mathbb{P}\{\zeta = k\}}{\mathbb{P}\{\zeta > 0\}} \\ &= \binom{J-1}{k} \left(\frac{1-\varrho}{\varrho}\right)^k \frac{(\varrho)^{J-1}}{1-(\varrho)^{J-1}}.\end{aligned}\quad (7)$$

The error probability for a subsequent transmission after the initial one when the PU is sensed idle, ϵ_n , is calculated considering the probability of not having other SUs transmitting for all possible combinations of SUs with packets to transmit.

$$\begin{aligned}\epsilon_n &= \sum_{k=1}^{J-1} \phi_k \left(1 - (1-p(1-P_{FA}))^k \times \right. \\ &\quad \left. (1-p(1-P_{FA})(1-\varrho))^{J-1-k}\right) \\ &= 1 - \frac{\varrho^{J-1} (1-p(1-P_{FA})(1-\varrho))^{J-1}}{1-\varrho^{J-1}} \times \\ &\quad \left(\left(1 + \left(\frac{1-\varrho}{\varrho}\right) \left(\frac{1-p(1-P_{FA})}{1-p(1-P_{FA})(1-\varrho)}\right)\right)^{J-1} - 1\right).\end{aligned}\quad (8)$$

The success probability for the further transmissions of a packet, ς_n , takes into account the SU's transmission probability,

$$\varsigma_n = p(1-P_{FA})(1-\epsilon_n). \quad (9)$$

B. Service Time

The expected duration of one irrelevant epoch (with the PU active or transitioning to a different state) is denoted by $\mathbb{E}[\delta_{PU}]$ and the expected duration of a continuous sequence of relevant frames (with an idle PU) is denoted by $\mathbb{E}[\delta_{\overline{PU}}]$, which are calculated as

$$\mathbb{E}[\delta_{PU}] = \sum_{l=0}^{\infty} (l+2) (\pi_{11})^l (1-\pi_{11}) = \frac{2-\pi_{11}}{1-\pi_{11}}, \quad (10)$$

and

$$\mathbb{E}[\delta_{\overline{PU}}] = \sum_{l=0}^{\infty} (l) (\pi_{00})^l (1-\pi_{00}) = \frac{\pi_{00}}{1-\pi_{00}}. \quad (11)$$

One SU waits an average number of slots equal to $\mathbb{E}[\Delta_{SU}^1]$, which may include several PU epochs, until reaching a relevant epoch,

$$\begin{aligned}\mathbb{E}[\Delta_{SU}^1] &= \pi_{00} \sum_{l=0}^{\infty} (l+1) (1-\pi_{00})^l (1+l\mathbb{E}[\delta_{PU}]) \\ &= \frac{2-\pi_{00}-\pi_{11}}{(1-\pi_{11})\pi_{00}} = 1 + \frac{\mathbb{E}[\delta_{PU}]}{\mathbb{E}[\delta_{\overline{PU}}]}.\end{aligned}\quad (12)$$

The expectation of the packet service time, $\mathbb{E}[\Delta_{SU}]$, is calculated using the average number of retransmissions needed until successfully transmitting the packet,

$$\begin{aligned}\mathbb{E}[\Delta_{SU}] &= \mathbb{E}[\Delta_{SU}^1] \varsigma_1 + \\ &\quad (1-\varsigma_1) \left(\sum_{k=0}^{\infty} (k+1) \mathbb{E}[\Delta_{SU}^1] \varsigma_n (1-\varsigma_n)^{k-1} \right) \\ &= \frac{1+\varsigma_n-\varsigma_1}{\varsigma_n} \mathbb{E}[\Delta_{SU}^1].\end{aligned}\quad (13)$$

C. Probability of Queue Empty

The probability of an SU having its queue empty is related with the network's utilization rate, determined using $\lambda \mathbb{E}[\Delta_{SU}]$, where λ denotes the average load per SU. For a M/G/1 queue with a time-invariant service time, Takács [6, pp. 66–76] shows that $\varrho = 1 - \lambda \mathbb{E}[\Delta_{SU}]$ for $\lambda \mathbb{E}[\Delta_{SU}] < 1$ and that $\varrho = 0$ for $\lambda \mathbb{E}[\Delta_{SU}] \geq 1$. The relation in (14) can be used to calculate ϱ for a given load λ using numerical calculation; it only has analytical solutions for small J values.

$$1-\varrho = \lambda \mathbb{E}[\Delta_{SU}]. \quad (14)$$

D. Goodput

The aggregated goodput of the J SUs, S_a , can be calculated considering the fraction of slots that effectively transmit useful information. It is given by

$$S_a = J(1-\varrho)(1-P_U) \frac{\mathbb{E}[\Delta_{SU}^1]}{\mathbb{E}[\Delta_{SU}]}, \quad (15)$$

where P_U denotes the probability of a slot belonging to a relevant epoch and the quotient accounts the slots lost due to packet retransmissions. P_U is equal to

$$P_U = \frac{\mathbb{E}[\delta_{PU}]}{\mathbb{E}[\delta_{PU}] + \mathbb{E}[\delta_{\overline{PU}}]} = \frac{(2-\pi_{11})(1-\pi_{00})}{2-\pi_{11}-\pi_{00}}. \quad (16)$$

Using (14) it is trivial that $S_a = J\lambda$ as long as the system is not saturated.

The maximum aggregated goodput S_{sat} , is obtained when the system is saturated, for $\varrho = 0$, with

$$S_{\text{sat}} = J(1-P_U)p(1-P_{FA})(1-p(1-P_{FA}))^{J-1}. \quad (17)$$

E. Delay

For Poisson sources, the individual transmission delay on a SU can be modelled by a M/G/1 queue [7]. The average system delay for a packet can be expressed as

$$\mathbb{E}[\Delta_T] = \mathbb{E}[\Delta_{SU}] + \frac{\lambda \mathbb{E}[(\Delta_{SU})^2]}{2(1-\lambda \mathbb{E}[\Delta_{SU}])} + \frac{\mathbb{E}[V]}{2}, \quad (18)$$

which decompose the total packet delay, $\mathbb{E}[\Delta_T]$, into the service time, the time waiting in the queue and the time waiting for the end of the current epoch.

The second moment of the service time, $\mathbb{E}[(\Delta_{SU})^2]$, is required to calculate the queueing delay. It is calculated estimating the variance of the service time, $\sigma_{\Delta_{SU}}^2$, using

$$\mathbb{E}[(\Delta_{SU})^2] = (\mathbb{E}[\Delta_T])^2 + \sigma_{\Delta_{SU}}^2. \quad (19)$$

Assuming that the SU's epochs are uncorrelated, the variance of the service time is the summation of the variances for each relevant slot, $\sigma_{\Delta_{SU}}^2$, where the node may transmit. On the other hand, this variance depends on the number of irrelevant epochs (21), which depend on the variance of the irrelevant epochs, $\sigma_{\delta_{PU}}^2$, calculated using (23).

$$\begin{aligned} \sigma_{\Delta_{SU}}^2 &= \sigma_{\Delta_{SU}}^2 \varsigma_1 + \\ & (1 - \varsigma_1) \left(\sum_{k=0}^{\infty} (k+1) \sigma_{\Delta_{SU}}^2 \varsigma_n (1 - \varsigma_n)^{k-1} \right) \\ &= \frac{1 + \varsigma_n - \varsigma_1}{\varsigma_n} \sigma_{\Delta_{SU}}^2, \end{aligned} \quad (20)$$

$$\sigma_{\Delta_{SU}}^2 = \pi_{00} \sum_{l=0}^{\infty} l (1 - \pi_{00})^l \sigma_{\delta_{PU}}^2 = \frac{1 - \pi_{00}}{\pi_{00}} \sigma_{\delta_{PU}}^2, \quad (21)$$

and

$$\begin{aligned} \sigma_{\delta_{PU}}^2 &= \mathbb{E}[(\delta_{PU})^2] - (\mathbb{E}[\delta_{PU}])^2 \\ &= \sum_{l=0}^{\infty} (l+2)^2 (\pi_{11})^l (1 - \pi_{11}) - (\mathbb{E}[\delta_{PU}])^2 \\ &= \frac{4 - 3\pi_{11} + (\pi_{11})^2}{(1 - \pi_{11})^2} - \frac{1}{(1 - \pi_{11})^2} = \frac{3 - 3\pi_{11} + (\pi_{11})^2}{(1 - \pi_{11})^2}. \end{aligned} \quad (22)$$

Finally, the expected time waiting for the end of the current epoch when a new packet arrives to the SU's queue is

$$\begin{aligned} \mathbb{E}[V] &= P_U \mathbb{E}[\delta_{PU}] + (1 - P_U) \\ &= 1 + \frac{(2 - \pi_{11})(1 - \pi_{00})}{(1 - \pi_{11})(2 - \pi_{11} - \pi_{00})}. \end{aligned} \quad (23)$$

IV. OPTIMIZATION

The performance of the protocol can be optimized by properly defining the p value when there is context awareness, namely about the number of SUs and on their average load.

Without any context awareness, p can be set with the value that maximizes the throughput, denoted by p_{sat}^* . It is calculated solving $\frac{\partial S_{\text{sat}}}{\partial p} = 0$. From

$$\frac{\partial S_{\text{sat}}}{\partial p} = \left(1 - \varrho - p \frac{\partial \varrho}{\partial p} \right) (1 - Jp(1 - P_{FA})(1 - \varrho)), \quad (24)$$

we get

$$p_{\text{sat}}^* (1 - \varrho) = \frac{1}{J(1 - P_{FA})}. \quad (25)$$

The maximum system's throughput, S_{sat}^* , is obtained for $p = p^*$,

$$S_{\text{sat}}^* = (1 - P_U) \left(1 - \frac{1}{J} \right)^{J-1}. \quad (26)$$

P_U	$\pi_{1,1}$	$\pi_{0,0}$	N_T
0.502905	0.936388	0.935747	425
N_S	γ	P_D	P_{FA}
42	77.815435	0.997827	0.005997

TABLE I
CONFIGURATION OF CR SYSTEMS USED IN SECTION. V-A

It can be easily proved that $\lim_{J \rightarrow \infty} S_{\text{sat}}^* = \frac{1 - P_U}{e}$, as expected.

When the SU's load and their number is known and the system is not saturated (i.e. $J\lambda < S_{\text{sat}}^*$), it is possible to define a context aware optimal p value, denoted by p^* , as the one which minimizes the average packet delay. p^* can be calculated solving $\frac{\partial \mathbb{E}[\Delta_T]}{\partial p} = 0$ with the constraint $S_a = \lambda J$. For the results presented in this paper, the derivative was calculated by definition, setting the increment to 10^{-7} , and the bisection method was used to find the solution to the equation.

V. PERFORMANCE ANALYSIS

This section presents a set of performance results for a CR with a single PU and a set of SUs with homogeneous Poisson traffic. The analytical model is validated using simulations, using a simulator implemented in Matlab, which also runs the sensing model presented in section II-B. The first subsection analyses the influence of p parameter and the second one how the system scales with an increased number of SUs.

A. Influence of p

The influence of p parameter was tested for $J = 2$ and $J = 3$ SUs, considering the PU and sensing parameters presented in table I, and simulations that covered 1202 SU frames.

Figure 3 depicts the value estimated for ϱ (also named as P_{QE}) using the analytical model and measured during the simulations. The systems are saturated for $\lambda = 0.13$ packets/frame, thus $\varrho = 0$. For $\lambda = 0.04$ packets/frame/SU and $\lambda = 0.07$ packets/frame/SU respectively for $J = 3$ SUs and $JU = 2$ SUs, the analytical value follows the simulation one for low p and near 1 values, but a deviation is visible for intermediate values. In this region, the approximation of considering the value constant is worst because the oscillation of the ϱ value over time is greater. However, it can be seen that the model can be used to estimate the p value that maximizes ϱ .

Figure 4 shows the throughput variation with p : it is null for $p = 0$ and $p = 1$ and has the maximum value defined by (25). As expected, the estimated throughput follows the measured one. As long as $S_{\text{sat}} > J\lambda$ the throughput is equal to load, and when the condition is not verified, the system saturates. The analytical model provides the exact p_{sat}^* value, which can be used to configure the SUs' p value when the number of active SUs can be estimated, but the load is unknown. If the load is known, a more precise optimal value can be provided, considering the p value that minimizes the delay in the intervals represented in figure 4 for $\lambda = 0.04$ packets/frame/SU and $\lambda = 0.07$ packets/frame/SU, where the throughput is equal to the offered load. The variation of the packet delay in these intervals is represented in figure

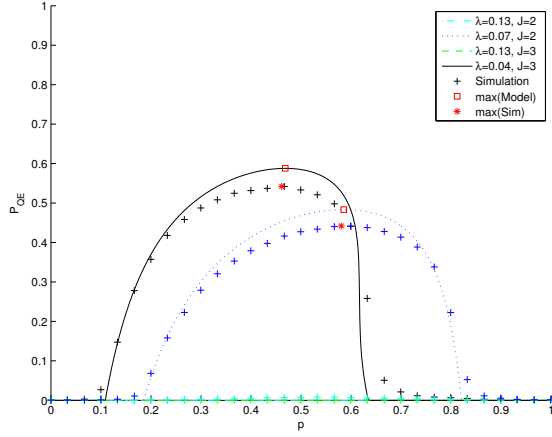


Fig. 3. ρ over p for different values of λ and J .

5, showing that the model can also be used to estimate the variation of the delay, although an offset is visible. Comparing the figures for ρ and delay, it can be shown that when the load is low, the minimum delay occurs for the p value that maximized ρ , thus providing an expedited way to calculate p^* for unsaturated systems with homogeneous load.

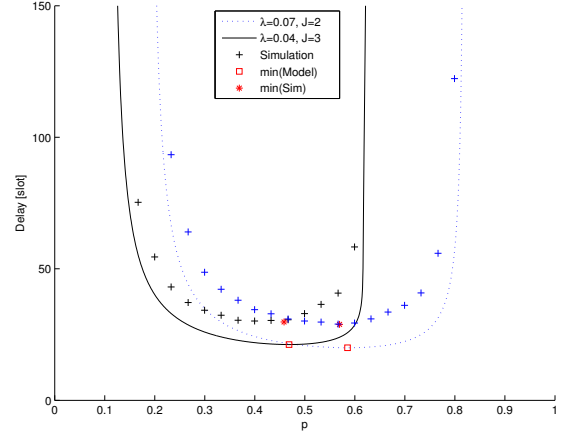


Fig. 5. Delay over p for different values of λ and J .

P_U	Simulation Time	$\pi_{1,1}$	$\pi_{0,0}$
0.502565	1052.539500	0.935762	0.935209
N_S	γ	P_D	P_{FA}
42	77.817634	0.997642	0.006173

TABLE II
CONFIGURATION OF CR SYSTEM USED IN SCALABILITY TESTS

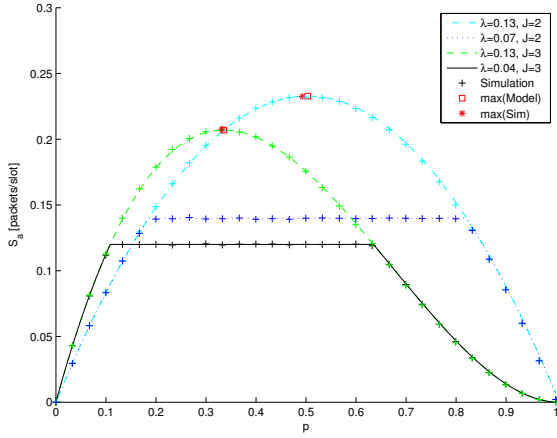


Fig. 4. Throughput over p for different values of λ and J .

B. Scalability

The scalability of the system with the number of SUs (J) was tested considering the PU and sensing parameters presented in table II, considering simulations with a duration of 1053 SU frames with $J = \{2, 3, 5, 7\}$ SUs. The p parameter was set to the value $p = 1/J \approx p_{sat}^*$, which only requires the knowledge about the number of active SUs. Figure 6 shows that ρ decreases faster with a load increase for a higher number of SUs. It also shows that the analytical model is slightly pessimistic in the determination of the saturation bound, since the saturation occurs for higher loads in the

simulation measurements. The throughput is depicted in figure

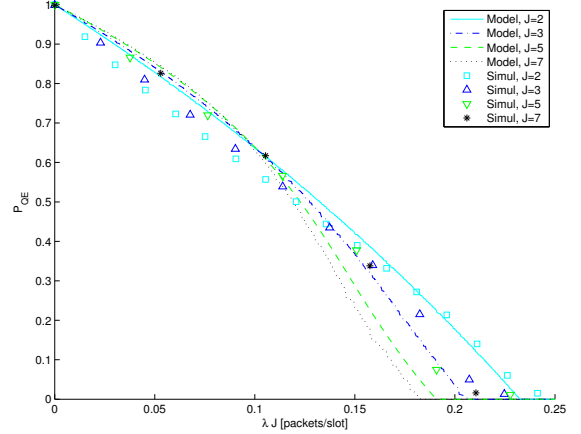


Fig. 6. ρ over λ for different values of J .

7, and shows that the network capacity decreases for a higher number of SUs, as expected from (17). The ρ deviation translate only into a slight deviation from the simulation results near the saturation limits. Figure 8 shows that a higher number of SUs leads to an higher packet transmission delay. Therefore, in order to handle higher loads or control delay, a channel access coordination mechanism must be implemented in the CR system to regulate the number of SUs accessing each channel.

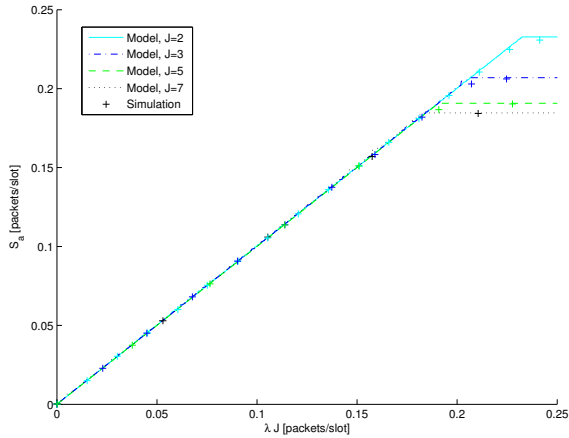


Fig. 7. Throughput over λ for different values of J .

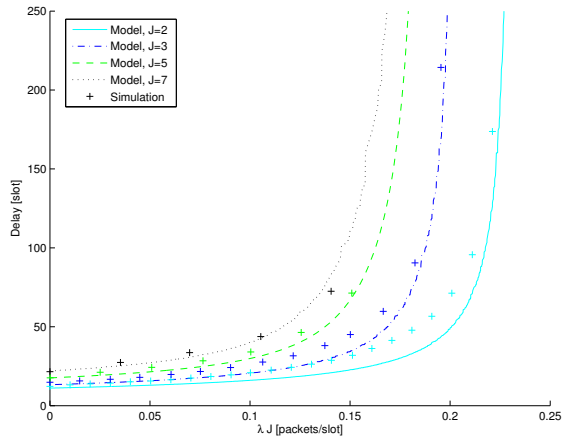


Fig. 8. Delay over λ for different values of J .

VI. CONCLUSION

This paper analysed the performance of a CR p -persistent Slotted Aloha protocol, where p can be set to p^* in a context aware system, calculated using the proposed model. It was shown that different context aware approaches can be followed, depending of what is known. When only the number of active BSs is known, it was shown that p_{sat}^* provides a good performance for a variable number of BSs. When the load is also known and is homogeneous, p^* can be calculated using the ϱ model. Therefore, it provides a mechanism to implement a context aware medium access control protocol.

Future work includes the development of efficient estimation mechanisms for the number of active SUs and for the average load generated by each one, which may be used to improve context awareness, and optimization mechanisms associated.

REFERENCES

- [1] Q. Chen, Y-C. Liang, M. Motani and W-C. Wong, "CR-CSMA: A Random Access MAC Protocol for Cognitive Radio Networks", in *Proc.*

- of *IEEE PIMRC*, pp. 486-490, Tokyo, Japan, Sept. 2009.
- [2] Q. Chen, M. Motani, W-C. Wong and Y-C. Liang, "Opportunistic Spectrum Access Protocol for Cognitive Radio Networks", in *Proc. of IEEE ICC*, Kyoto, Japan, June 2011.
- [3] Q. Chen, W-C. Wong, M. Motani and Y-C. Liang, "MAC Protocol Design and Performance Analysis for Random Access Cognitive Radio Networks", *IEEE J. Sel. Areas Commun.*, Vol. 31, pp. 2289-2300, Nov. 2013.
- [4] S. Hu, Y-D. Yao and A.U. Sheikh, "Slotted aloha for cognitive radio users and its tagged user analysis", in *Proc. of 21st Annual Wireless and Optical Communications Conference (WOCC)*, Kaohsiung, China, Apr. 2012.
- [5] M. Luis, A. Furtado, R. Oliveira, R. Dinis and L. Bernardo, "Towards a Realistic Primary Users Behavior in Single Transceiver Cognitive Networks", *IEEE Communications Letters*, Vol. 17, pp. 309-312, Feb. 2013.
- [6] L. Takács, *Introduction to the Theory of Queues*, Oxford University Press, 1962.
- [7] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed., Prentice-Hall, 1992.

Bibliografia

- [BG92] D. P. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, 2nd edition, Jan. 1992.
- [BGdMT06] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. Wiley-Interscience, 2nd edition, May 2006.
- [Bia00] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, March 2000.
- [CB05] D. Cabric and R.W. Brodersen. Physical layer design issues unique to cognitive radio systems. *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2:759–763, 2005.
- [CC09] C. Cormio and K. R. Chowdhryt. A Survey On MAC Protocols For Cognitive Radio Networks. *Ad Hoc Networks* 7, pages 1315–1329, Feb. 2009.
- [CLMW09] Q. Chen, Y.-C. Liang, M. Motani, and W.-C. Wong. CR-CSMA: A Random Access MAC Protocol for Cognitive Radio Networks. *2009 IEEE 20th International Symposium on Indoor and Mobile Radio Communications*, pages 486–490, 2009.
- [CLMW11] Q. Chen, Y.-C. Liang, M. Motani, and W.-C. Wong. Opportunistic Spectrum Access Protocol for Cognitive Radio Networks. *2011 IEEE International Conference on Communications (ICC)*, pages 1–6, June 2011.
- [CLMW13] Q. Chen, Y.-C. Liang, M. Motani, and W.-C. Wong. MAC Protocol Design and Performance Analysis for Random Access Cognitive Radio Networks. *IEEE Journal on Selected Areas in Communications*, 31(11):2289–2300, Nov 2013.
- [CTB06] D. Cabric, A. Tkachenko, and R. Brodersen. Spectrum sensing measurements of pilot, energy and collaborative detection. In *Proc. IEEE Military Communications Conference*, pages 1–7, D.C, USA, Oct. 2006.

- [Dai05] J. N. Daigle. The Basic M/G/1 Queueing System. *Queueing Theory with Applications to Packet Telecommunication*, pages 159–223, 2005.
- [DSB12] A. Domenico, E. Strinati, and M. Benedetto. A survey on MAC Strategies for Cognitive Radio Networks. *IEEE Communications Surveys and Tutorials*, 14(1):21–44, 2012.
- [FGK13a] A. B. Flores, R. E. Guerra, and E. W. Knightly. IEEE 802.11af: A Standard for TV White Space Spectrum Sharing. *IEEE Communications Magazine*, 14(10):92–100, October 2013.
- [FGK⁺13b] A. B. Flores, R. E. Guerra, E. W. Knightly, P. Ecclesine, and S. Pandey. IEEE 802.11af: A Standard for TV White Space Spectrum Sharing. *IEEE Communications Magazine*, 51(10):92–100, Oct. 2013.
- [Gar91] W. A. Gardner. Exploitation of spectral redundancy in cyclostationary signals. *IEEE Signal Processing Magazine*, 8(2):14–36, 1991.
- [GBE⁺08] G. Gurney, G. Buchward, L. Ecklund, S. Kuffner, and J. Grosspietsch. Geo-location Database Techniques For Incumbent Protection in the TV White Space. *3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2008 DySPAN*, pages 14–17, 2008.
- [GS07] A. Ghasemi and Elvino S. Sousa. Optimization of spectrum sensing for opportunistic spectrum access in cognitive radio networks. *IEEE 4th Consumer Communications and Networking Conference (CCNC 2007)*, pages 1022–1026, Jan. 2007.
- [HDML08] K.D. Huang, K.R. Duffy, D. Malone, and D.J. Leith. Investigating the validity of IEEE 802.11 MAC modeling hypotheses. *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008*, pages 1–6, Sept. 2008.
- [HYS12] S. Hu, Y-D. Yao, and A. U. Sheikh. Slotted Aloha for Cognitive Radio Users and its Tagged User Analysis. In *2012 21st Annual Wireless and Optical Communications Conference (WOCC)*, pages 1–5, 2012.
- [Kle76] L. Kleirock. *Queueing systems volume 2: Computer applications*. Wiley-Interscience, Jan. 1976.
- [LFO⁺12] M. Luis, A. Furtado, R. Oliveira, R. Dinis, and L. Bernardo. Energy Sensing Parameterization Criteria for Cognitive Radios. *International Symposium on Wireless Communication Systems (ISWCS 2012)*, pages 61–65, Aug. 2012.

- [LFO⁺13] M. Luis, A. Furtado, R. Oliveira, R. Dinis, and L. Bernardo. Towards a Realistic Primary Users Behavior in Single Transceiver Cognitive Networks. *IEEE Communications Letters*, 17(2):309–312, Feb 2013.
- [LG08] J. Little and S. C. Graves. Little’s Law. *International Series in Operations Research and Management Science 115*, page 81, March 2008.
- [LKHP07] J. Lunden, V. Koivunen, A. Huttunen, and H. V. Poor. Spectrum sensing in cognitive radios based on multiple cyclic frequencies. *2nd International Conference on Cognitive Radio Oriented Wireless Network and Communications (CrownCom 2007)*, pages 37–43, 2007.
- [MBA⁺07] K. Maeda, A. Benjebbour, T. Asai, T. Furuno, and T. Ohya. Recognition among OFDM-based systems utilizing cyclostationarity-inducing transmission. *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pages 516–523, 2007.
- [MDV⁺01] M. Mehta, N. Drew, G. Vardoulas, N. Greco, and C. Niedermeier. Reconfigurable terminals: an overview of architectural solutions. *IEEE Communications Magazine*, 39(8):82–89, 2001.
- [MMB07] S. M. Mishra, R. Mahadevappa, and R. W. Brodersen. Cognitive technology for ultra-wideband/WiMax coexistence. In *2nd IEEE International Symposium on New Frontiers In Dynamic Spectrum Access Networks*, pages 179–186, Dublin, Ireland, Apr. 2007.
- [MSR07] L. Ma, C.-C. Shen, and B. Ryu. A survey on MAC Strategies for Cognitive Radio Networks. *Proceedings of IEEE DySPAN*, pages 547–558, April 2007.
- [OK09] R. Oliveira and I. Koutsopoulos. Queue and Channel State Awareness for Maximum Throughput Access Control in CSMA/CA-Based Wireless LANs. *IEEE Wireless Communications and Networking Conference (WCNC 2009)*, pages 1–6, April 2009.
- [SCL⁺09a] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell. IEEE 802.22 Working Group on Wireless Regional Area Networks. *IEEE Communications Magazine*, 47(1):130–138, Jan. 2009.
- [SCL⁺09b] C. R. Stevenson, G. Chouinard, Z. Lei, S. Shellhammer, and W. Caldwell. IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard. *IEEE Communications Magazine*, 47(1):130–138, Jan. 2009.

- [SKMK10] K. G. Shin, H. Kim, A. W. Min, and A. Kumar. Cognitive Radios For Dynamic Spectrum Access: From Concept To Reality. *IEEE Wireless Communications*, 17(6):64–74, 2010.
- [SW94] B. Shawyer and B. Watson. *Borel’s Methods of Summability: Theory and Application*. Oxford University Press, Oct. 1994.
- [SZ12] H. Su and X. Zhang. Opportunistic MAC protocols for cognitive radio based wireless networks. *41st Annual Conference on Information Sciences and Systems*, pages 363–368, March 2012.
- [Tak62] L. Takács. Introduction to the Theory of Queues. *Oxford University Press*, 35(1):450–453, 1962.
- [Tan03] A. S. Tanenbaum. *Computer Networks*. Prentice Hall PTR, 4th edition, 2003.
- [TCB07] A. Tkachenko, D. Cabric, and R. W. Brodersen. Cyclostationary Feature Detector Experiments Using Reconfigurable BEE2. In *2nd IEEE International Symposium on New Frontiers In Dynamic Spectrum Access Networks (DySPAN 2007)*, pages 216–219, Dublin, Ireland, Apr. 2007.
- [Tur04] G. L. Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 2004.
- [Urk67] H. Urkowitz. Energy Detection of Unknown Deterministic Signals. *Proceedings of the IEEE*, 55(4):523–531, 1967.
- [VFECH01] G. Vardoulas, J. Faroughi-Esfahani, G. Clemo, and R. Haines. Blind Radio Access Technology Discovery and Monitoring for Software Define Radio Communication Systems: Problems and Techniques. pages 306–310, London, U.K, 2001. Second International Conference on 3G Mobile Communication Technologies.
- [YA06] T. Yucek and H. Arslan. Spectrum characterization for opportunistic cognitive radio systems. pages 1–6, Washington, D.C., USA, Oct. 2006. IEEE Military Communications Conference (MILCOM 2006).
- [YA09] T. Yucek and H. Arslan. A survey of Spectrum Sensing Algorithms for Cognitive Radio Applications. *IEEE Communications Surveys and Tutorials*, 11(1):116–130, 2009.
- [ZC08] C. Zhou and C. Chigan. A Game Theoretic DSA-Driven MAC Framework for Cognitive Radio Networks. *IEEE International Conference on Communications (ICC 2008)*, pages 4165–4169, May 2008.



